PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/705, G01N 33/68, C12Q 1/68, C07K 16/28

A2

(11) International Publication Number: WO 99/38975

(43) International Publication Date: 5 August 1999 (05.08.99)

(21) International Application Number: PCT/US99/02033

(22) International Filing Date: 29 January 1999 (29.01.99)

(30) Priority Data:
60/073,190 30 January 1998 (30.01.98) US

(71) Applicant *(for all designated States except US)*: UNIVERSITY TECHNOLOGY CORPORATION [US/US]; Suite 250, 3101 Iris Avenue, Boulder, CO 80301 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: JOHNSON, Thomas, E. [US/US]; 1121 West Enclave Circle, Louisville, CO 80027 (US). SIKELA, James, M. [US/US]; 6046 S. Kingston Circle, Englewood, CO 80111 (US). SIMPSON, Victoria, J. [US/US]; 1121 W. Enclave Circle, Englewood, CO 80027 (US). RIKKE, Brad, A. [US/US]; 750 Hartford, Boulder, CO 80303 (US).

(74) Agents: POLIZZI, Catherine, M. et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES ASSOCIATED WITH CNS DEPRESSANT SENSITIVITY AND METHODS OF USE THEREOF

(57) Abstract

The invention provides polynucleotide and polypeptide sequences associated with CNS depressant sensitivity, and methods for identifying and classifying candidate CNS depressants as well as methods to identify agents which may modulate CNS depressant action. The invention also provides mouse mGluR5 polynucleotide and polypeptide sequences.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES ASSOCIATED WITH CNS DEPRESSANT SENSITIVITY AND METHODS OF USE THEREOF

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the priority benefit of U.S. Provisional Patent Application No. 60/073,190, filed January 30, 1998. The priority application is hereby incorporated herein by reference in its entirety.

STATEMENT OF RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH

This invention was made in part during work supported by grants from the National Institutes of Health (NIH) NIAAA RO1-AA08940 and NIGMS RO1-GM55635. The government has certain rights in this invention.

TECHNICAL FIELD

This invention relates to polynucleotide and polypeptide sequences associated with CNS depressant sensitivity, as well as methods using these sequences.

BACKGROUND ART

Central nervous system (CNS) depressants have been used for a variety of purposes, including anesthesia. CNS depressants are also socially, economically and medically significant in terms of drug use and abuse, such as ethanol, barbituate, and sedative consumption. While several useful anesthetics have been developed, such as propofol, improved anesthetics are still needed. Furthermore, the basis of propensity toward excess ethanol consumption, or differential sensitivity toward this and other CNS depressants, is still not known.

Several animal models have been used to study the basis of CNS depressant action. Simpson and Johnson (1996) *Int. Rev. Neurobiol.* 39:223-241. LS (long sleep) and SS (short sleep) mice were initially selected on the basis of differential ethanol sensitivity. Simpson and Johnson (1996); McClearn and Kakihana (1981) "Development of Animal Models as Pharmacogentic Tools", McClearn et al, eds, Research Monograph No. 6 at 147-159 (U.S. Dept. Health Human Services). In addition to ethanol, the LS and SS mice are

differentially sensitive to a variety of other agents having properties of CNS depressants. Simpson et al. (1996) *Anesth. Analg.* 82:1-5; Markel et al. (1996) *Mam. Genome* 7:408-412; Simpson and Johnson (1996). These CNS depressants include chloral hydrate, enflurane, etomidate, flurazepam, halogenated ethanols, isoflurane, ketamine, midazolam, paraldehyde, MK-801, urethane and propofol. These agents have diverse chemical structures suggesting that the LS and SS mice have been selected, at least in part, for genes that influence sensitivity to multiple classes of CNS depressants. However, LS and SS mice do not differ in sensitivity to ether or halothane, (Baker et al. (1980) *Pharmacol. Biochem. and Behav.* 12:691-695), indicating that there has not been uniform selection for differential sensitivity to all anesthetic agents.

LS and SS mice exhibit significant CNS differences in their sensitivity to propofol-induced anesthesia. Simpson and Blednov (1996) *Anesth. Analg.* 82:327-331. Given the same 20 mg/kg dose via the retroorbital sinus, SS mice slept about 3.5 minutes (SEM = 0.4 min) and LS mice slept about 2.2 minutes longer, with no difference between males and females. Moreover, although propofol has a potent effect on $GABA_A$ receptors, (Tanelian et al. (1993) *Anesthesiology* 78:757-776), LS and SS mice did not exhibit significant differences in $GABA_A$-activated chloride channels. However, the genetic locus contributing to this differential sensitivity has not been discovered.

The metabotropic glutamate receptors (mGluRs) are G protein-coupled receptors capable of activating a variety of intracellular second messenger systems following the binding of glutamate or other potent agonists including quisqualate and 1-aminocyclopentane-1,3-dicarboxylic acid (*trans*-ACPD). Schoepp et al. (1990) *Trends Pharmacol. Sci.* 11:508; Schoepp and Conn (1993) *Trends Pharmacol. Sci.* 14:13.

Activation of different metabotropic glutamate receptor subtypes *in situ* elicits one or more of the following responses: activation of phospholipase C, increases in phosphoinositide (PI) hydrolysis, intracellular calcium release, activation of phospholipase D, activation or inhibition of adenylyl cyclase, increases and decreases in the formation of cyclic adenosine monophosphate (cAMP), activation of guanylyl cyclase, increases in the formation of cyclic guanosine monophosphate (cGMP), activation of phospholipase A_2, increases in arachidonic acid release, and increases or decreases in the activity of voltage- and ligand-gated ion channels (Schoepp and Conn (1993); Schoepp (1994) *Neurochem. Int.* 24:439; Pin and Duvoisin (1995) *Neuropharmacology* 34:1).

Thus far, eight distinct mGluR subtypes have been isolated via molecular cloning, and named mGluRl to mGluR8 according to the order in which they were discovered. See, for example, Nakanishi (1994) *Neuron* 13:1031; Pin and Duvoisin (1995) *Neuropharmacology* 34:1; Knopfel et al. (1995) J. *Med. Chem.*. 38:1417. Further diversity occurs through the expression of alternatively spliced forms of certain mGluR subtypes. Pin et al. (1992) *Proc. Natl. Acad. Sci.* 89:10331; Minakami et al. (1994) *BBRC* 199:1136. All of the mGluRs appear to be structurally similar, in that they are proposed to be single subunit membrane proteins possessing a large amino-terminal extracellular domain (ECD) followed by seven putative transmembrane domain (7TMD) comprising seven putative membrane spanning helices connected by three intracellular and three extracellular loops, and an intracellular carboxy-terminal domain of variable length (cytoplasmic tail) (CT). The eight mGluR subtypes have unique patterns of expression within the mammalian CNS that in many instances are overlapping. Masu et al. (1991) *Nature* 349:760; Martin et al. (1992) *Neuron* 9:259; Ohishi et al. (1993) *Neurosci,* 53:1009; Tanabe et al. (1993) *J. Neurosci.* 13:1372; Ohishi et al. (1994) *Neuron* 13:55; Abe et al. (1992) *J. Biol. Chem.* 267:13361; Nakajima et al. (1993) *J. Biol. Chem.* 267:13361; Nakajima et al. (1993) *J. Biol. Chem.* 268:11868; Okamoto et al. (1994) *J. Biol. Chem.*269:1231; Duvoisin et al. (1995) *J. Neurosci.* 15:3075.

The eight mGluRs have been subdivided into three groups based on amino acid sequence identities, the second messenger systems they utilize, and pharmacological characteristics. Nakanishi (1994) *Neuron* 13:1031; Pine and Duvoisin (1995) *Neuropharmacology* 34:1; Knopfel et al. (1995) *J. Med. Chem.* 38:1417.

The Group I mGluRs comprise mGluRl, mGluR5 and their alternatively spliced variants. The binding of agonists to these receptors results in the activation of phospholipase C and the subsequent mobilization of intracellular calcium. For example, *Xenopus* oocytes expressing recombinant mGluRl receptors have been utilized to demonstrate this effect indirectly by electrophysiological means. Masu et al. (1991); Pin et al. (1992). Similar results were achieved with oocytes expressing recombinant mGluR5 receptors. Abe et al. (1992); Minakami et al. (1994). Alternatively, agonist activation of recombinant mGluRl receptors expressed in Chinese hamster ovary (CHO) cells stimulated PI hydrolysis, cAMP formation, and arachidonic acid release as measured by standard biochemical assays. Aramori and Nakanishi (1992) *Neuron* 8:757. In comparison,

activation of mGluR5 receptors expressed in CHO cells stimulated PI hydrolysis and subsequent intracellular calcium transients but no stimulation of cAMP formation or arachidonic acid release was observed. Abe et al (1991).

The Group II mGluRs include mGluR2 and mGluR3. Activation of these receptors as expressed in CHO cells inhibits adenylyl cyclase activity via the inhibitory G protein, G_i, in a pertussis toxin-sensitive fashion. Tanabe et al. (1992) *Neuron* 8:169; Tanabe et al. (1993).

The Group III mGluRs include mGluR4, mGluR6, mGluR7 and mGluR8. Like the Group II receptors these mGluRs are negatively coupled to adenylate cyclase to inhibit intracellular cAMP accumulation in a pertussis toxin-sensitive fashion when expressed in CHO cells (Tanabe et al. (1993); Nakajima et al.(1993); Okamoto et al. (1994); Duvoisin et al. (1995).

Several reports indicate that various mGluRs may be affected by various anesthetics. See, e.g., Dildy-Mayfield et al. (1996) *J. Pharmacol. Exp. Ther.* 276:1058-1065 (describing experiments addressing the effect of various anesthetics on certain mGluRs, but not mGluR5); Orser et al. (1995) *Br. J. Pharmacol.* 116:1761-1768 (describing experiments addressing inhibition by propofol of N-methyl-D-aspartate subtype of glutamate receptor); Miyamoto et al. (1994) *Eur. J. Pharmacol.* 260:99-102 (describing experiments addressing delay in recorery from halothane anesthesis by agonists for metabotropic glutamate receptors).

Nucleic acids encoding human metabotropic glutamate receptors mGluR1, mGluR2, mGluR3 and mGluR5 are described in U.S. Pat. 5,521,297. See also U.S. Pat. 5,202,257. Human mGluR5 is cloned and characterized by Minakami et al. (1994) *Biochem. Biophys. Res. Comm.* 199:1136-1143. Rat mGluRs polynucleotide and polypeptide sequences are described in Abe et al. (1992). See also Minakami et al. (1993) *Biochem. Biophys. Res. Comm.* 194:622-627. The 550 base nucleotide sequence containing a single exon (of about 100 nucleotides) of mouse mGluR5 has been reported. Minakami et al. (1995) *J. Neurochem.* 65:1536-1542.

There is a continuing need to develop new CNS depressants, particularly anesthetics, which exert fewer side effects. A defined molecular target may allow thedesign of anesthetics which act specifically at neural tissue, thus avioiding side effects related to non-specific actions at other organ systems. There is also a need to refine and/or

modulate the action of anesthetics, particularly in certain individuals, such as at-risk individuals.

All references cited herein are incorporated by reference in their entirety.

DISCLOSURE OF THE INVENTION

The present invention provides methods using polynucleotide and polypeptide sequences associated with CNS sensitivity, as well as the polynucleotide and polypeptide sequences themselves. In particular, these methods are used to identify, inter alia, agents which may exhibit CNS depressant activity and/or modulate CNS depressant activity.

Accordingly, in one aspect, the invention provides a method for identifying an agent which may exhibit CNS depressant activity, comprising the steps of: (a) introducing a polynucleotide comprising a mouse polynucleotide sequence associated with CNS depressant sensitivity into a suitable host cell, wherein the mouse polynucleotide sequence corresponds to a polynucleotide sequence of yeast artificial chromosome YRT2; (b) contacting host cell of step (a) with at least one agent to be tested; and (c) analyzing at least one characteristic associated with expression of the mouse YRT2 polynucleotide, wherein an agent is identified by its ability to modulate expression of the mouse YRT2 polynucleotide.

In another aspect, the invention provides a method of classifying a CNS depressant, comprising the steps of: (a) introducing a polynucleotide comprising a mouse polynucleotide sequence associated with CNS depressant sensitivity into a suitable host cell, wherein the mouse polynucleotide sequence corresponds to a polynucleotide sequence of yeast artificial chromosome YRT2; (b) contacting host cell of step (a) with at least one agent to be tested; (c) analyzing at least one characteristic associated with expression of the mouse YRT2 polynucleotide, wherein an agent is identified by its ability to modulate expression of the mouse YRT2 polynucleotide, and wherein a CNS depressant which modulates expression may fall within a class of CNS depressants which displays differential reactivity in LS and SS mice.

In another aspect, the invention provides a method of identifying an agent which may modulate CNS depressant sensitivity in an individual, said method comprising: (a) introducing a polynucleotide comprising a mouse polynucleotide sequence associated with CNS depressant sensitivity into a suitable host cell, wherein the mouse polynucleotide

sequence corresponds to a polynucleotide sequence of yeast artificial chromosome YRT2; (b) contacting host cell of (a) with at least one agent to be tested; and (c) analyzing at least one characteristic associated with expression of the mouse YRT2 polynucleotide, wherein an agent is identified by its ability to modulate expression of the mouse YRT2 polynucleotide.

In another aspect, the invention provides a method of identifying an agent which may exhibit CNS depressant activity, comprising the steps of: (a) introducing a polynucleotide comprising a metabotropic glutamate receptor polynucleotide sequence into a suitable host cell; (b) contacting host cell of step (a) with at least one agent to be tested; (c) analyze at least one characteristic associated with expression of the metabotropic glutamate receptor polynucleotide, wherein an agent is identified by its ability to modulate expression of the mammalian metabotropic glutamate receptor polynucleotide. The metabotropic glutamate receptor may be mammalian.

In another aspect, the invention provides a method of classifying a CNS depressant, comprising the steps of: (a) introducing a polynucleotide comprising a mammalian metabotropic glutamate receptor polynucleotide sequence into a suitable host cell; (b) contacting host cell of step (a) with at least one agent to be tested; (c) analyzing at least one characteristic associated with expression of the metabotropic glutamate receptor polynucleotide, wherein an agent is identified by its ability to modulate expression of the metabotropic glutamate receptor polynucleotide sequence, wherein a CNS depressant which modulates the metabotropic glutamate receptor polynucleotide expression may fall within a class of CNS depressants which displays differential reactivity in LS and SS mice.

In another aspect, the invention provides an isolated polynucleotide comprising a polynucleotide encoding a mouse mGluR5 polypeptide, wherein the mouse mGluR5 polypeptide is at least 10 contiguous amino acids of SEQ ID NO:2 and exhibits mGluR5 activity, and wherein the at least 10 contiguous amino acids are not depicted in SEQ ID NO:6 or SEQ ID NO:7.

In another aspect, the invention provides an isolated polynucleotide comprising a polynucleotide of at least about 25 contiguous nucleotides of SEQ ID NO:1, wherein the at least about 25 contiguous nucleotides are not depicted in SEQ ID NO:3 or SEQ ID NO:5.

In another aspect, the invention provides a polynucleotide comprising a region of at least 25 contiguous nucleic acids of SEQ ID NO:1, said region having at least about 97%

sequence identity to a sequence in SEQ ID NO:1. The region is not depicted in SEQ NO:3 or SEQ ID NO:5.

In another aspect, the invention provides an isolated polypeptide comprising at least about 5 contiguous amino acids of the sequence of SEQ ID NO:2, wherein the isolated polypeptide exhibits an mGluR5 function, and wherein the about 5 contiguous amino acids are not depicted in SEQ ID NO:6 or SEQ ID NO:7.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a dose-response curve depicting propofol dose (milligrams per kilogram) versus sleep time (minutes) in Long Sleep (LS; open triangle) and Short Sleep (SS; solid triangle) mouse lines. Sleep time is defined as the duration of loss of righting reflex. Each point represents the mean sleep time ± SEM for n = 6-10. For every dose tested, the LS and SS lines are differentially sensitive ($P < 0.01$).

Fig. 2 is a bar graph depicting sleep times (minutes) in LS (open bars) and SS (hatched bars) mice. Routes of administration were jugular cannula (JUG; left bar of each pair) or retroorbital sinus (ROS; right bar of each pair) injections. Each bar represents means sleep time ± SEM (n = 7-9). There was a small, statistically significant difference in Short Sleep (SS) sleep times between the two techniques ($P < 0.02$) but no difference in Long Sleep (LS) sleep times ($P = 0.09$).

Fig. 3 is a bar graph depicting effects of propofol on muscimol-stimulated $^{36}Cl^-$ uptake in Long Sleep (LS, open bars) and Short Sleep (SS, hatched bars) brain membranes. Bars express the mean ± SEM. Eight to 10 female LS or SS mice were used per group.

Figs. 4 (A) and (B) are graphs depicting duration of LORR (loss of righting reflex) and BLA (brain levels of propofol at awakening) in pigmented and albino LSXSS RI strains. Pigmented strains are shown as black bars; albino RI strains are shown as white bars. Error bars indicate standard error of the mean. In Fig. 4 (B), the strain distribution pattern mean LORR for propofol for the pigmented RI strains is indicated by the filled ovals, and the means for the albino RI strains is indicated by the open ovals. The numbers next to the ovals indicate the RI strain(s) having that mean. RI strain 32 had a mean LORR of 4.4 minutes (SEM = 0.7) for females and 6.6 minutes (SEM = 1.1) for males, but it was not included in this analysis because at that time it was still segregating for the albino mutation.

Fig. 5 depicts a genetic map of murine chromosome 7 showing LOD scores for positioning *Lorpl* using the LSXSS RIs. Markers include the albino mutation (c), the *pink-eyed dilution* mutation (p) and six *D7Mit* SSLPs. The dashed line indicates the recommended threshold for statistical significance. Data were combined for males and females.

Fig. 6 depicts a genetic map of murine chromosome 7 showing confirmation of *Lorpl* from F_2 intercrosses between ILS and ISS mice. Markers are *D7Mit* SSLPs. Marker *D7Mit31* is within the *Tyr* gene near the site of albino mutation, *c*. The dashed line indicates the recommended threshold for statistical significance. Data were combined for males and females (N = 164).

Figs. 7 (A) and (B) are bar graphs depicting propofol LORRs of (ISS *c/c x* C57BL/6 c^{2J}/C) F_1s and of ILS and ISS mice with tyrosine. ILS (black bars) and ISS (white bars) mice coinjected with tyrosine (indicated as pr + tyr), substrate of the tyrosinase enzyme encoded by the albino gene, also did not exhibit significant differences in LORR ($p = 0.78$, ANOVA two-tailed significance) from controls at a concentration of tyrosine (200 mg/kg) previously shown to differentially alter ethanol-induced LORRs in LS and SS. "n" indicates the number of mice (males and females) tested. Error bars indicate standard error of the mean.

Fig. 8 depicts a genetic map of murine chromosome 7 showing LOD scores for positioning *Lorpl* using the LSXSS RIs. Results are shown for propofol, isoflurane, enflurane, and ethanol. The left-hand panels show data for females; the right-hand panels show data for males.

Fig. 9 depicts a genetic map of murine chromosome 7 showing LOD scores for positioning *Lorpl* using LSXSS RIs. This experiment was performed using etomidate.

Fig. 10 depicts a genetic map of murine chromosome 7 showing confirmation of *Lorpl* from F_2 intercrosses between ILS and ISS mice. This experiment was performed using etomidate.

Fig. 11 is a schematic depiction of YAC YRT2, containing a 250 kb genomic insert containing the mouse tyrosinase and mGluR5 genes. Expanded depiction of vector arms are shown below. Markers and functional elements from the vector arms are shown as filled arrows (except the open circle, which indicates the centromere). Numbers 1-5 indicate the exons of the tyrosinase gene. Tel, telomeres; Amp, ampicillin resistance gene;

Tk, thymidine kinase gene of herpes simplex virus; Cen, centromere (*CEN4*); associated with the Gal-1 promoter (*GAL1*); *TRP1* and *URA3*, yeast markers.

Fig. 12 is a bar graph comparing sleep times of SS (non-transgenic; left bar of each pair) and transgenic mice containing YRT2 (right bar of each pair) when given propofol. N = number of injections.

Fig. 13 is a schematic showing the strategy for analyzing mouse DNA sequences present on YRT2.

Fig. 14 (A)-(G) depicts the nucleotide sequence and a conceptual amino acid translation of most of the mouse mGluR5 coding region (SEQ ID NOS:1 and 2).

Fig. 15 shows a portion of mouse mGluR5 nucleotide sequence obtained from the Genbank database and its conceptual amino acid translation (SEQ ID NOS:3 and 4).

Fig. 16 (A)-(G) depicts the nucleotide and amino acid sequences of rat mGluR5 (SEQ ID NOS:5 and 6).

Fig. 17 (A)-(C) depicts the amino acid sequence of human mGluR5 (SEQ ID NO:7).

MODES FOR CARRYING OUT THE INVENTION

We have discovered mouse polynucleotide sequences that are associated with a particular, significant, biological phenomenon, namely sensitivity to CNS depressants. In particular, the sensitivity referred to herein (which may also be considered as decreased resistance) is with respect to a particular class of CNS depressants defined by the differential reactivity of LS and SS mice to these CNS depressants. This class of CNS depressants includes, but is not limited to, ethanol, chloral hydrate, enflurane, etomidate, flurazepam, halogenated ethanols, isoflurane, ketamine, midazolam, paraldehyde, MK-801, urethane and propofol. This discovery makes possible useful screening assays for agents which may exhibit CNS depressant activity as well as agents which may modulate CNS depressant sensitivity. This discovery may also lead to finding corresponding human sequences which confer differential sensitivity to CNS depressants, including ethanol, which in turn could provide the basis for powerful diagnostics in the context of, for example, ethanol sensitivity or vulnerability to ethanol addiction as well as determination of presence of risk factors for anesthesia administration.

Mouse polynucleotide sequences were identified based on association with a particular biological function, namely, resistance to CNS depressant activity. Briefly, the polynucleotide and polypeptide sequences that form the basis of the screening methods described herein were discovered and obtained by: (a) mapping the genetic locus associated with the phenomenon of differential CNS depressant response; (b) complementation of the SS (resistant) phenotype to an LS (sensitive) phenotype by a YAC denoted YRT2 containing a 250 kb mouse genomic insert containing a tightly linked locus to the observed differential response phenomenon, namely, the *albino* (*tyr*) locus encoding tyrosinase. Thus, sequences contained in this YAC are associated with differential CNS depressant response. A schematic of YRT2 is depicted in Fig. 11.

The mapping experiments were conducted with propofol, enflurane, isoflurane, ethanol, and etomidate. In all cases, the trait associated with the differential response (the *Lorpl* locus) was significantly linked to the murine tyrosinase (albino) locus on chromosome 7. Thus, the sequences associated this differential sensitivity likely encompass the entire class of CNS depressants described above.

Definitions

As used herein, a "YRT2 polynucleotide" is a mouse polynucleotide sequence that is contained in, or corresponds to, a polynucleotide sequence within the yeast artificial chromosome (YAC) vector denoted YRT2, which is described and discussed herein. A schematic depiction of YRT2 is provided in Fig. 11. As the mouse polynucleotide sequences reflect a genomic configuration (i.e., the insert was obtained from a mouse genomic library), and the identity of at least part of the YAC insert is known, as well as the insert's genomic location, it is clear to those skilled in the art that these sequences may be obtained from sources other than YRT2 itself. It is understood that the definition of "YRT2 polynucleotide" includes, but is not limited to, polynucleotides contained within the tyrosinase gene and/or the mGluR5 gene. A gene, as understood by those of skill in the art, contains coding as well as non-coding sequences (such as 5' and 3' flanking sequences and introns). Accordingly, any definitions pertaining to YRT2 polynucleotides also apply to tyrosinase and mGluR5 polynucleotides.

A "YRT polypeptide" is a polypeptide that is encoded by a YRT2 polynucleotide. It is understood that the definition of "YRT2 polypeptide" includes, but is not limited to,

polypeptides encoded within the tyrosinase gene and/or the mGluR5 gene. Accordingly, any definitions pertaining to YRT2 polypeptides also apply to tyrosinase and mGluR5 polypeptides.

A sequence "corresponding" to a sequence in YRT2 means that, when compared (i.e., aligned), the sequences share at least about 80% sequence identity, preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, even more preferably at least about 95% sequence identity, even more preferably at least about 98% sequence identity, most preferably 100% sequence identity. As discussed below, the term "corresponding" is used to emphasize that the sequences used in the methods described in this invention need not literally have been obtained from YRT2, but need only correspond in terms of contiguous nucleic acid composition to the sequences of YRT2. It is also understood that a sequence corresponding to YRT2 also applies to any polypeptide sequences encoded in the mouse polynucleotide sequence of YRT2, and that, as noted above for polynucleotide sequence identity, need not be a perfect correspondence to the sequence of YRT2 and need not be obtained by expressing YRT2 sequences per se.

As used herein, "expression" encompasses any level leading to, involving, and/or resulting from production of a protein product. Accordingly, expression includes any aspect of transcription (including the modulation or effect of control sequences such as enhancers or promoters), translation, processing, sorting, post-translational modification(s), folding, conformation, assembly, subcellular and/or extracellular localization and/or position, binding, and/or effector function(s) (such as those involved in a signal transduction, or second messenger, pathway).

A characteristic which is associated with "modulation" of expression is a characteristic which is associated with an alteration, increase or decrease, in expression.

A "control region" of a gene is any sequence, of any length, which affects gene expression, most usually transcription. Examples of control regions include, but are not limited to, promoters and enhancers.

An "mGluR polynucleotide" refers to a polynucleotide contained within (or in) an mGluR gene. Examples of mGluR genes have been discussed and are known in the art. The mGluR polynucleotide may be of any origin, preferably eukaryotic, more preferably mammalian, such as mouse, human, or rat. However, it is understood that, due to sequence homology, an mGluR polynucleotide may also be obtained from other cell types, such as C.

elegans, Drosophila, and yeast. For example, GenBank provides a homolog of mGluR as derived from C. elegans.

As used herein, "mGluR5 polynucleotide" refers to a polynucleotide contained within (or in) an mGluR5 gene. A "mouse mGluR5 polynucleotide" refers to a polynucleotide contained within (or in) a mouse mGluR5 gene. Similarly, a "mammalian mGluR5 polynucleotide" or "human mGluR5 polynucleotide" refers to a mammalian or human polynucleotide contained within or in a mammalian or human mGluR5 gene, respectively. A sequence of part of the mouse mGluR5 gene is depicted within SEQ ID NO:1. A "fragment" or "region" of mGluR5 gene is a portion of the mGluR5 gene, and as such may contain coding and/or non-coding sequences. For example, SEQ ID NO:1 contains a fragment of the mGluR5 in depicting the coding sequence. Preferably, a fragment of mGluR5 comprises at least 10 contiguous nucleotides, more preferably at least 15, more preferably at least 25, more preferably at least 30, more preferably at least 50, more preferably at least 100, more preferably at least 150, more preferably at least 200, more preferably at least 250, more preferably at least 300 contiguous nucleotides.

"mGluR5 " refers to a protein (polypeptide) product encoded in a mGluR5 gene. As noted for mGluR5 polypeptides above, a type (indicating genetic origin) of mGluR5 is specified, such as mouse, human, or mammalian. SEQ ID NO:2 depicts a conceptual translation of an open reading frame of mouse mGluR5. A "fragment" or "region" of mGluR5 (i.e., full-length) is a portion of the mGluR5 gene product. It is understood that mGluR5 may exist in more than one form, due to, for example, alternative splicing.

"mGluR5 function" or "mGluR5 activity" refers to an activity or characteristic associated with expression of mGluR5 . These functions include, but are not limited to, (a) transcription; (b) translation, including post-translational processing; (c) regulation by other polypeptides; (d) binding ligand; (e) signal transduction pathway (i.e., second messenger system) activity. For mGluR5, these signal transduction pathway activities include, but are not limited to, PI hydrolysis and mobilization of intracellular calcium.

As used herein, a characteristic which is associated with a "modulation" of mGluR5 function or activity is a characteristic which is associated with an alteration, increase or decrease, in mGluR5 function or activity. As used herein, "modulation of mGluR5 activity" means any change in any measurable parameter along the functional pathway of mGluR5. The "functional pathway" of mGluR5 includes transcription, translation,

processing (including pre-and post-translational processing), positioning and/or conformation of receptor (either in solution or within a membrane context), binding to receptor, activities triggered by binding to the receptor (i.e., signal transduction pathway), and/or modulation of other neual receptor activities.

"Suitable reaction conditions" refer to those conditions which allow a characteristic to be measured to occur. For example, if the characteristic is binding, then suitable reaction conditions are those which allow this binding to occur. If the characteristic is a signal transduction activity, then suitable reaction conditions are those which allow this activity to occur. Depending on which characteristic will be measured, the reaction conditions will vary. Determination of suitable reaction conditions is well within the skill of those in the art.

A sequence, whether polynucleotide or polypeptide, "depicted in" a SEQ ID NO, means that the sequence is present as an identical contiguous sequence in the sequence of the SEQ ID NO. Conversely, a contiguous sequence that is "not depicted in" a SEQ ID NO means that the contiguous sequence is not present as an identical contiguous sequence in the sequence of the SEQ ID NO. The term "contiguous" sequence and referring to, for example, an amino acid sequence of a specified length, are interchangeable in this context.

As used herein, a "polynucleotide" is a polymeric form of nucleotides of any length, which contain deoxyribonucleotides, ribonucleotides, and/or their analogs. The terms "polynucleotide" and "nucleotide" as used herein are used interchangeably. Polynucleotides may have any three-dimensional structure, and may perform any function, known or unknown. The term "polynucleotide" includes double- , single-stranded, and triple-helical molecules. Unless otherwise specified or required, any embodiment of the invention described herein that is a polynucleotide encompasses both the double-stranded form and each of two complementary single-stranded forms known or predicted to make up the double stranded form. Not all linkages in a polynucleotide need be identical.

The following are non-limiting examples of polynucleotides: a gene or gene fragment, exons, introns, mRNA, tRNA, rRNA, ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, nucleic acid probes, primers, and adaptors. A polynucleotide may comprise modified nucleotides, such as methylated nucleotides and

nucleotide analogs. The use of uracil as a substitute for thymine in a deoxyribonucleic acid is also considered an analogous form of pyrimidine.

In the context of polynucleotides, a "linear sequence" or a "sequence" is an order of nucleotides in a polynucleotide in a 5' to 3' direction in which residues that neighbor each other in the sequence are contiguous in the primary structure of the polynucleotide. A "partial sequence" is a linear sequence a portion of a polynucleotide, wherein the polynucleotide is known to comprise additional residues in one or both directions.

If present, modification to the nucleotide structure may be imparted before or after assembly of the polymer. The sequence of nucleotides may be interrupted by non-nucleotide components. A polynucleotide may be further modified after polymerization, such as by conjugation with a labeling component. Other types of modifications included in this definition are, for example, "caps", substitution of one or more of the naturally occurring nucleotides with an analog, internucleotide modifications such as, for example, those with uncharged linkages (e.g., methyl phosphonates, phosphotriesters, phosphoamidates, cabamates, etc.) and with charged linkages (e.g., phosphorothioates, phosphorodithioates, etc.), those containing pendant moieties, such as, for example, proteins (e.g., nucleases, toxins, antibodies, signal peptides, poly-L-lysine, etc.), those with intercalators (e.g., acridine, psoralen, etc.), those containing chelators (e.g., metals, radioactive metals, boron, oxidative metals, etc.), those containing alkylators, those with modified linkages (e.g., α-anomeric nucleic acids, etc.), as well as unmodified forms of the polynucleotide(s).

Further, any of the hydroxyl groups ordinarily present in the sugars may be replaced by phosphonate groups, phosphate groups, protected by standard protecting groups, or activated to prepare additional linkages to additional nucleotides, or may be conjugated to solid supports. The 5' and 3' terminal OH groups can be phosphorylated or substituted with amines or organic capping group moieties of from 1 to 20 carbon atoms. Other hydroxyls may also be derivatized to standard protecting groups.

Polynucleotides can also contain analogous forms of ribose or deoxyribose sugars that are generally known in the art, including, but not limited to, 2'-O-methyl-, 2'-O-allyl, 2'-fluoro- or 2'-azido-ribose, carbocyclic sugar analogs, α-anomeric sugars, epimeric sugars such as arabinose, xyloses or lyxoses, pyranose sugars, furanose sugars, sedoheptuloses, acyclic analogs and abasic nucleoside analogs such as methyl riboside.

Although conventional sugars and bases will be used in applying the method of the invention, substitution of analogous forms of sugars, purines and pyrimidines can be advantageous in designing a final product, as can alternative backbone structures like a polyamide backbone.

A polynucleotide or polynucleotide region has a certain percentage (for example, 75%, 80%, 85%, 90%, 95%, 98%, or 99%) of "sequence identity" to another sequence means that, when aligned, that percentage of bases are the same in comparing the two sequences. As is discussed further herein, "sequence identity" can be indicated by one of several measurable parameters, such as sequence alignment techniques (most typically performed with the use of standard alignment programs) and hybridization techniques, both of which are discussed herein.

"Hybridization" refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson-Crick base pairing, Hoogstein binding, or in any other sequence-specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi-stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of a PCR, or the enzymatic cleavage of a polynucleotide by a ribozyme.

When hybridization occurs in an antiparallel configuration between two single-stranded polynucleotides, those polynucleotides are described as "complementary". A double-stranded polynucleotide can be "complementary" to another polynucleotide if hybridization can occur between one of the strands of the first polynucleotide and the second. Complementarity (the degree that one polynucleotide is complementary with another) is quantifiable in terms of the proportion of bases in opposing strands that are expected to form hydrogen bonding with each other, according to generally accepted base-pairing rules.

A "primer" is a short polynucleotide, generally with a free 3' -OH group, that binds to a target potentially present in a sample of interest by hybridizing with the target, and thereafter promoting polymerization of a polynucleotide complementary to the target.

An "adaptor" is a short, partially-duplexed polynucleotide that has a blunt, double-stranded end and a protruding, single-stranded end. It can be ligated, through its double-stranded end, to the double-stranded end of another polynucleotide. This provides known

sequences at the ends of thus modified polynucleotides. Often adaptors contain specific sequences for primer binding and/or restriction endonuclease digestion.

A "probe" when used in the context of polynucleotide manipulation refers to a polynucleotide which is provided as a reagent to detect a target potentially present in a sample of interest by hybridizing with the target. Usually, a probe will comprise a label or a means by which a label can be attached, either before or subsequent to the hybridization reaction. Suitable labels include, but are not limited to radioisotopes, fluorochromes, chemiluminescent compounds, dyes, and enzymes.

"Transformation" or "transfection" refers to the insertion of an exogenous polynucleotide into a host cell, irrespective of the method used for the insertion, for example, lipofection, transduction, infection or electroporation. The exogenous polynucleotide may be maintained as a non-integrated vector, for example, a plasmid, or alternatively, may be integrated into the host cell genome.

A polynucleotide is said to "encode" a polypeptide if, in its native state or when manipulated by methods well known to those skilled in the art, it can be transcribed and/or translated to produce the polypeptide or a fragment thereof. For purposes of this invention, and to avoid cumbersome referrals to complementary strands, the anti-sense (or complementary) strand of such a polynucleotide is also said to encode the sequence; that is, a polynucleotide sequence that "encodes" a polypeptide includes both the conventional coding strand and the complementary sequence (or strand).

The terms "polypeptide", "oligopeptide", "peptide" and "protein" are used interchangeably herein to refer to polymers of amino acids of any length. The polymer may be linear or branched, it may comprise modified amino acids, it may be interrupted by non-amino acids, and it may be assembled into a complex of more than one polypeptide chain. The terms also encompass an amino acid polymer that has been modified naturally or by intervention; for example, disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, or any other manipulation or modification, such as conjugation with a labeling component. Also included within the definition are, for example, polypeptides containing one or more analogs of an amino acid (including, for example, unnatural amino acids, etc.), as well as other modifications known in the art.

In the context of polypeptides, a "linear sequence" or a "sequence" is an order of amino acids in a polypeptide in an N-terminal to C-terminal direction in which residues that

neighbor each other in the sequence are contiguous in the primary structure of the polypeptide. A "partial sequence" is a linear sequence of a portion of a polypeptide, wherein the polypeptide is known to comprise additional residues in one or both directions.

A polypeptide "fragment" (also called a "region") is a polypeptide comprising an amino acid sequence that has at least 5 contiguous amino acids of a sequence, more preferably at least 10 contiguous amino acids, more preferably at least about 15 contiguous amino acids, even more preferably at least about 25 contiguous amino acids, even more preferably at least about 30 contiguous amino acids, even more preferably at least about 40 contiguous amino acids, or even larger.

A polypeptide or polypeptide region (or fragment) has a certain percentage of "sequence identity" (for example, 75%, 80%, 85%, 90%, 95%, 98%, 99%) to another sequence means that, when aligned, that percentage of amino acids are the same in comparing the two sequences. As discussed herein, those skilled in the art generally use commercially available alignment programs to determine degree of sequence identity.

A "fusion polypeptide" is a polypeptide comprising regions in a different position than occurs in nature. The regions may normally exist in separate proteins and are brought together in the fusion polypeptide, or they may normally exist in the same protein but are placed in a new arrangement in the fusion polypeptide.

A "functionally preserved" variant of a YRT2 polynucleotide or YRT2 polypeptide is a sequence which retains at least one aspect of YRT2 function. Functionally preserved variants can be due to differences in linear sequence, arising from, for example, single base mutation(s), addition(s), deletion(s), and/or modification(s) of the bases. The difference can also arise from changes in the sugar(s) and/or linkage(s) between the bases. Regarding polypeptides, functionally preserved variants may arise, for example, by conservative and/or non-conservative amino acid substitutions, amino acid analogs, and deletions. The function that is preserved depends upon the relevant function being considered. For example, a YRT2 polynucleotide is considered for its ability to encode a YRT2 polypeptide (or fragment thereof), then the ability of a variant sequence to encode the same polypeptide is the relevant function. If a YRT2 polypeptide is considered for its ability to bind to a particular entity (such as an antibody or ligand), then the ability of a variant sequence to encode a polypeptide with equivalent binding characteristics that is relevant.

"Recombinant," as applied to a polynucleotide or gene, means that the polynucleotide is the product of various combinations of cloning, restriction and/or ligation steps, and other procedures that result in a construct that is distinct from a polynucleotide found in nature.

A "vector" is a self-replicating nucleic acid molecule that transfers an inserted nucleic acid molecule into and/or between host cells. The term includes vectors that function primarily for insertion of a nucleic acid molecule into a cell, replication of vectors that function primarily for the replication of nucleic acid, and expression vectors that function for transcription and/or translation of the DNA or RNA. Also included are vectors that provide more than one of the above functions.

"Expression vectors" are defined as polynucleotides which, when introduced into an appropriate host cell, can be transcribed and translated into a polypeptide(s). An expression vector also comprises control elements operatively linked to the encoding region to enable and/or facilitate expression of the polypeptide in the target cell. An "expression system" usually connotes a suitable host cell comprised of an expression vector that can function to yield a desired expression product.

A "host cell" includes an individual cell or cell culture which can be or has been a recipient for vector(s) or for incorporation of nucleic acid molecules and/or proteins. Host cells include progeny of a single host cell, and the progeny may not necessarily be completely identical (in morphology or in genomic of total DNA complement) to the original parent cell due to natural, accidental, or deliberate mutation. A host cell includes cells transfected in vivo with a polynucleotide(s) of this invention.

A "cell line" or "cell culture" denotes eukaryotic cells, derived from higher, multicellular organisms, grown or maintained in vitro. It is understood that the descendants of a cell may not be completely identical (either morphologically, genotypically, or phenotypically) to the parent cell. Cells described as "uncultured" are obtained directly from a living organism, and are generally maintained for a limited amount of time away from the organism (i.e., not long enough or under conditions for the cells to undergo substantial replication).

"Heterologous" means derived from (i.e., obtained from) a genotypically distinct entity from the entity to which it is being compared. For example, a polynucleotide may be placed by genetic engineering techniques into a plasmid or vector derived from a different

source, thus becoming a heterologous polynucleotide. A promoter which is linked to a coding sequence with which it is not naturally linked is a heterologous promoter.

An "isolated" or "purified" polynucleotide, polypeptide, antibody or cell is one that is substantially free of the materials with which it is associated in nature. By substantially free is meant at least 50%, preferably at least 70%, more preferably at least 80%, and even more preferably at least 90% free of the materials with which it is associated in nature. As used herein, an "isolated" polynucleotide or polypeptide also refers to recombinant polynucleotides or polypeptides, which, by virtue of origin or manipulation: (1) are not associated with all or a portion of a polynucleotide or polypeptide with which it is associated in nature, (2) are linked to a polynucleotide or polypeptide other than that to which it is linked in nature, or (3) does not occur in nature, or (4) in the case of polypeptides arise from expression of recombinant polynucleotides. Thus, for example, an isolated substance may be prepared by using a purification technique to enrich it from a source mixture. Enrichment can be measured on an absolute basis, such as weight per volume of solution, or it can be measured in relation to a second, potentially interfering substance present in the source mixture. Increasing enrichments of the embodiments of this invention are increasingly more preferred. Thus, for example, a 2-fold enrichment is preferred, 10-fold enrichment is more preferred, 100-fold enrichment is more preferred, 1000-fold enrichment is even more preferred. A substance can also be provided in an isolated state by a process of artificial assembly, such as by chemical synthesis or recombinant expression.

A "stable duplex" of polynucleotides, or a "stable complex" formed between any two or more components in a biochemical reaction, refers to a duplex or complex that is sufficiently long-lasting to persist between formation of the duplex or complex and subsequent detection, including any optional washing steps or other manipulation that may take place in the interim.

A substance is said to be "selective" or "specific" if it reacts or associates more frequently, more rapidly, with greater duration and/or with greater affinity with a particular cell or substance than it does with alternative cells or substances. An antibody "specifically binds" to a target if it binds with greater affinity, avidity, more readily, and/or with greater duration than it binds to other substances.

As used herein, the term "agent" means a biological or chemical compound such as a simple or complex organic or inorganic molecule, a peptide, a protein or an oligonucleotide. A vast array of compounds can be synthesized, for example oligomers, such as oligopeptides and oligonucleotides, and synthetic organic compounds based on various core structures, and these are also included in the term "agent". In addition, various natural sources can provide compounds for screening, such as plant or animal extracts, and the like. Compounds can be tested singly or in combination with one another.

A "polymerase chain reaction" ("PCR") is a reaction in which replicate copies are made of a target polynucleotide using one or more primers, and a catalyst of polymerization, such as a reverse transcriptase or a DNA polymerase, and particularly a thermally stable polymerase enzyme. Methods for PCR are taught in U.S. Patent Nos. 4,683,195 (Mullis) and 4,683,202 (Mullis et al.). All processes of producing replicate copies of the same polynucleotide, such as PCR or gene cloning, are collectively referred to herein as "replication."

An "antibody" (interchangeably used in plural form) is an immunoglobulin molecule capable of specific binding to a target, such as a polypeptide, through at least one antigen recognition site, located in the variable region of the immunoglobulin molecule. An antibody can be from any source of animal capable of producing them, for example, mouse, rat, rabbit, or human antibodies. As used herein, the term encompasses not only intact antibodies, but also fragments thereof (such as Fab, Fab', $F(ab')_2$, Fv, single chain (ScFv)), mutants thereof, fusion proteins, humanized antibodies, and any other modified configuration of the immunoglobulin molecule that comprises an antigen recognition site of the required specificity. The term "antibody" includes polyclonal antibodies and monoclonal antibodies.

"Immunological recognition" or "immunological reactivity" refers to the specific binding of a target through at least one antigen recognition site in an immunoglobulin or a related molecule, such as a B cell receptor or a T cell receptor.

The term "antigen" refers to the target molecule that is specifically bound by an antibody through its antigen recognition site. The antigen may, but need not be chemically related to the immunogen that stimulated production of the antibody. The antigen may be polyvalent, or it may be a monovalent hapten. Examples of kinds of antigens that can be

recognized by antibodies include polypeptides, polynucleotides, other antibody molecules, oligosaccharides, complex lipids, drugs, and chemicals.

An "immunogen" is an antigen capable of stimulating production of an antibody when injected into a suitable host, usually a mammal. Compounds may be rendered immunogenic by many techniques known in the art, including crosslinking or conjugating with a carrier to increase valency, mixing with a mitogen to increase the immune response, and combining with an adjuvant to enhance presentation.

General Techniques

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology (including recombinant techniques), microbiology, cell biology, biochemistry and immunology, which are within the skill of the art. Such techniques are explained fully in the literature, such as: "Molecular Cloning: A Laboratory Manual", second edition (Sambrook et al., 1989); "Oligonucleotide Synthesis" (M.J. Gait, ed., 1984); "Animal Cell Culture" (R.I. Freshney, ed., 1987); "Methods in Enzymology" (Academic Press, Inc.); "Handbook of Experimental Immunology" (D.M. Wei & C.C. Blackwell, eds.); "Gene Transfer Vectors for Mammalian Cells" (J.M. Miller & M.P. Calos, eds., 1987); "Current Protocols in Molecular Biology" (F.M. Ausubel et al., eds., 1987); "PCR: The Polymerase Chain Reaction", (Mullis et al., eds., 1994); "Current Protocols in Immunology" (J.E. Coligan et al., eds., 1991); Manipulating the Mouse Embryo, second edition (Hogan et al., 1994); "Modern Biological Analysis of an Organism" (D.C. Shakes, eds., 1995).

Screening methods

The present invention encompasses a variety of screening methods using the polynucleotide and/or polypeptide sequences described herein. These methods may be used as a basis for classifying agents in any of a number of ways, including, but not limited to, (a) agents which may exert CNS depressant activity, including candidate anesthetics; (b) agents that may be included in the class of agents in response to which LS and SS mice display differential sensitivity; (c) agents that may be included in the class of agents in response to which there is no differential sensitivity between LS and SS mice. The agents identified and classified under (b) and/or (c) may represent agents which act under two

different kinds of mechanisms. Screening methods of this invention may also be used to identify agents which may modulate CNS depressant activity. Such an agent may be useful, for example, in administering a CNS depressant such as an anesthetic to certain at-risk or overly resistant individuals. Such an agent may also find use in administration to individuals who are overly sensitive to ethanol consumption or who are at risk of developing an addiction to ethanol.

These methods may be practiced in a variety of embodiments. The methods described herein encompass in vitro, cell-based, and in vivo screening assays. In the in vitro embodiments, an agent is tested for its ability to modulate function in a system which does not use intact cells. In the cell-based embodiments, living cells containing sequences described herein are used for testing agents. In the in vivo embodiments, transgenic mice harboring sequences described herein are used for testing agents.

Some of these screening embodiments employ mouse polynucleotide or polypeptide sequences contained in (or encoded in) YAC YRT2, or sequences corresponding to those sequences contained in (or encoded in) YAC YRT2. It is not necessary to know the identity of the sequence(s) (i.e., the individual contiguous nucleotide and/or amino acid identities) to practice these methods. For example, the sequence may be contained in a fragment (such as a restriction fragment or a fragment obtained by PCR) of YRT2, without knowledge of the nucleotide sequence identity of the fragment. Alternatively, a particular known sequence of YRT2 (or known sequence corresponding to a sequence of YRT2) may be used. It is not necessary to obtain sequences for these methods from YRT2 per se, as long as the sequence(s) correspond to those contained in (or encoded in) YRT2.

YRT2 is a YAC containing a 250 kb mouse genomic insert is depicted in Fig. 11 and is described by Schedl et al. (1993) *Nature* 362:258-261. YRT2 was obtained from a YAC library of C3H mouse DNA. Robertson (1987) "Tetracarcinomas and Embryonic Stem Cells" (Robertson, ed.) at 71-112. YRT2 contains an 80kb polynucleotide sequence corresponding to the tyrosinase gene (Fig. 11). We have discovered that the remaining 170 kb contains the gene encoding mGluR5, a metabotropic glutamate receptor (discussed below). There may be other coding and/or control sequences in YRT2. Accordingly, another way of describing the sequences of YRT2, any portion of which may be used in the methods of this invention, are those mouse genomic sequences which encompass the region

corresponding to the entire tyrosinase gene and another approximately 170 kb upstream of the tyrosinase gene.

How generally to prepare any of these sequences or relevant fragments is discussed in a later section addressing polynucleotides and polypeptides of the invention. It is understood that the description of preparation of sequences applies to sequences to be used for the screening methods described herein. Because the genomic location of the mouse sequences of YRT2 are known, and sequences contained within YRT2 are known, one skilled in the art could obtain sequences corresponding to sequences contained within YRT2 from sources other than YRT2 by using, for example, PCR methods applied to a suitable source of mouse DNA (i.e., a source of mouse DNA that would contain the desired sequences), such as commercially available libraries, genomic preparations, cDNA preparations, and DNA obtained by excised gel fragments. Alternatively, because we have shown that the *Lorpl* locus is tightly linked to the tyrosinase gene, these sequences could be used to obtain those and surrounding sequences from a murine genomic (or cDNA) library. Accordingly, mGluR5 sequences (discussed below) and/or tyrosinase sequences may also be used to obtain the desired sequences for these screening methods.

It is also understood that the sequences used in these screening methods can also include regulatory (non-coding) sequences, such as enhancers and/or promoters. When these sequences are used, reporter systems indicating regulatory region function may be employed. Such systems are well-known in the art, and include, but are not limited to, luciferase; aequorian (i.e., green fluorescent protein from *Aequorea victoria*); β-galactosidase; chloramphenicol acetyl transferase; immunologically detectable protein "tags" such as human growth hormone; and the like. See, for example, Current Protocols in Molecular Biology (F.M. Ausubel et al., eds., 1987) and periodic updates. Any assay which detects a product of the reporter gene, either by directly detecting the protein encoded by the reporter gene or by detecting an enzymatic product of a reporter gene-encoded enzyme, is suitable for use in the present invention. Assays include colorimetric, fluorimetric, or luminescent assays or even, in the case of protein tags, radioimmunoassays or other immunological assays. Many of these assays are commercially available.

As noted in the definition of "agent" provided above, the agent may be any compound, complex or substance. Generally, the choice of agents to be screened is governed by several parameters, such as the particular polynucleotide or polypeptide target,

its perceived function, its three-dimensional structure (if known or surmised), and other aspects of rational drug design. Techniques of combinatorial chemistry can also be used to generate numerous permutations of candidates. Those of skill in the art can devise and/or obtain suitable agents for testing.

An agent is generally identified by its ability to modulate expression of the polynucleotide. As discussed, above, modulation of expression of a polynucleotide may occur at any level that affects its function. An agent may modulate polynucleotide expression by preventing, reducing or increasing transcription. An example of such an agent is one that binds to the upstream controlling region, including a polynucleotide sequence or polypeptide. An agent may modulate expression by preventing, reducing, or increasing translation of its corresponding mRNA. An example of such an agent is one that binds to the mRNA, such as an anti-sense polynucleotide, or an agent which selectively degrades the mRNA, or an agent that selectively stabilizing the mRNA. An agent may modulate expression by binding to the expressed polypeptide. An example of such an agent is a polypeptide or a chelator. Examples of the effect of such binding agents may include the degradation of the polypeptide, increased half-life of the polypeptide, prevention of polypeptide interaction with a ligand, and the stabilization of the polypeptide with a ligand.

In some embodiments, sequences contained in an mGluR5 gene (including 5' and 3' flanking (non-coding) sequences and introns), which encodes mGluR5, a subclass of metabotropic glutamate receptors, are used. We have discovered evidence that mouse mGluR5 sequences may be involved in this differential response to CNS depressants. For example, a comparison in which LS and SS mice were treated with an agonist of Group I mGluR (which includes mGluR5), ACPD, resulted in a greater change in increased resistance in SS (resistant) mice when compared to LS mice (Example 5). In another experiment, transgenic mice harboring the YAC in an SS background exhibited significantly longer sleep times when treated with an Group I agonist than control (non-transgenic) mice receiving the same agonist (Example 5). These observations suggest that a pathway(s) involving mGluR5 function may play a role in CNS sensitivity (and/or resistance). However, because of the complexity of CNS depressant reaction, and multiple pathways, the inventors are not confining their theory of the sequences associated with the differential CNS depressant response to mGluR5 gene or gene product. Other sequences

on the YAC YRT2 described herein may be associated with this phenomenon, and are included in the screening methods described herein.

Further, because we have observed that mGluR5 appears to play a role in CNS depressant sensitivity in mice, we believe that this receptor plays an analogous functional role on other mammalian species, including human. The high degree of homology among mouse, rat, and human mGluR5 further supports this belief. See SEQ ID NOS:5 and 6, which depict polynucleotide encoding rat mGluR5 (96% overall homology according to a BLAST search comparison) and the amino acid sequence of rat mGluR5 (over 96% overall homology (sequence identity), according to a BLAST search comparison); see also SEQ ID NO:7, which depicts human mGluR5 (over 95% overall homology (sequence identity), according to a BLAST search comparison). Accordingly, the screening methods herein employ mGluR5 sequences, such as mammalian mGluR5, including, but not limited to, mouse, rat, and human. It is also possible, if not likely, that other mGluRs play a part in CNS depressant response. Accordingly, the methods described herein also employ mGluR sequences other than mGluR5, particularly those receptors in class I (mGluR1 and mGluR5 and alternatively spliced variants) and class II (mGluR2 and mGluR3 and alternatively spliced variants). The other mGluR sequences that may be used in these methods include, not only mammalian sequences, but also non-mammalian sequences which show significant (i.e., over 50%) sequence identity to mammalian mGluR sequences. Examples of sources of these other mGluR sequences include, but are not limited to, C. elegans (nemotode) and Drosophila. These sequences are readily available to those skilled in the art, given their publication on public databases such as Genbank and routine methods for obtaining and/or synthesizing polynucleotide sequences.

For any of these mGluR sequences, non-coding as well as coding regions may be used. For example, upstream control regions, such as promoters and/or enhancers, could be obtained and tested. If a non-coding control region is not publicly available on a database such as Genbank, these sequences may be obtained by using routine techniques in the art, such as chromosome walking using publicly available sequences as probes. For example, a cDNA probe containing coding sequences for an mGluR, such as human mGluR5, could be used to probe a human genomic library (or a genomic or cDNA library of another organism, if homologous sequences are desired). Positive, overlapping clones are then analyzed by, for example, restriction analysis. If sequences are desired even further from

the probe sequence, sequentially overlapping clones may be obtained using probes derived from previously identified positive clones, thus effecting walking.

It is also understood that the screening methods of this invention include structural, or rational, drug design, in which the amino acid sequence, three-dimensional atomic structure or other property (or properties) of YRT2 polynucleotide or polypeptide (or mGluR5 polynucleotide or polypeptide) provides a basis for designing an agent which is expected to bind to YRT2 polynucleotide or polypeptide (or mGluR5 polynucleotide or polypeptide). Generally, the design and/or choice of agents in this context is governed by several parameters, such as the perceived function of the YRT2 polynucleotide or polypeptide (or mGluR5 polynucleotide or polypeptide) target, its three-dimensional structure (if known or surmised), and other aspects of rational drug design. Techniques of combinatorial chemistry can also be used to generate numerous permutations of candidate agents.

The screening methods described above represent primary screens, designed, inter alia, to detect any agent that may exhibit CNS depressant activity, or modulate such activity. The skilled artisan will recognize that secondary tests will likely be necessary in order to evaluate an agent further. For example, because a characteristic associated with expression of a YRT2 polynucleotide is generally assessed, it may be desirable to further confirm that this characteristic is involved with CNS depressant activity. One way of accomplishing this would be to conduct the experiment using a known CNS depressant. If the response (i.e., the impact on expression of the YRT2 polynucleotide or mGluR5 polynucleotide) is similar or the same as that response caused by the agent to be tested, it is reasonable to surmise that the characteristic observed is involved in CNS depressant activity.

In vitro embodiments

In in vitro screening methods of this invention, an agent is screened in an in vitro system, which may include either of the following: (1) an assay for an agent which modulates the translation of a YRT2 (or mGluR5) mRNA or a YRT2 polynucleotide (or an mGluR polynucleotide, including mGluR5) encoding a polypeptide; (2) an assay for an agent that binds to a YRT2 (or mGluR5) polynucleotide(s) or polypeptides (or to a mammalian mGluR polypeptide, including mGluR5).

Accordingly, the invention includes methods of identifying an agent which modulates a YRT2 polynucleotide encoding a YRT2 polypeptide, comprising combining at least one agent to be tested with a YRT2 polynucleotide which encodes a YRT2 polypeptide under suitable reaction conditions, and determining whether the at least one agent modulates this polynucleotide by analyzing at least one characteristic associated with translation of the YRT2 polynucleotide. The invention also includes methods of identifying an agent which binds to a YRT2 polynucleotide or polypeptide, comprising combining at least one agent with a YRT2 polynucleotide or polypeptide under suitable reaction conditions, and determining whether the at least one agent binds to the YRT2 polynucleotide or YRT2 polypeptide. Examples of these types of assays are provided below. These methods also pertain to using mGluR polypeptide(s) and/or polynucleotide(s), such as mammalian mGluR, such as mGluR5, which includes mouse mGluR5. For the discussion which follows, it is understood that the YRT2 sequences serve as example and that these techniques pertain also to mGluR sequences.

For an assay that determines whether an agent modulates the translation of YRT2 mRNA or a polynucleotide encoding a YRT2 polypeptide, an in vitro transcription/translation system may be used. These systems are available commercially and provide an in vitro means to produce mRNA corresponding to a polynucleotide sequence of interest. After YRT2 mRNA is made, it can be translated in vitro and the translation products compared. Comparison of translation products between an in vitro expression system that does not contain any agent (negative control) with an in vitro expression system that does contain an agent indicates whether the agent is affecting translation. Comparison of translation products between control and YRT2 polynucleotides indicates whether the agent, if acting on this level, is selectively affecting translation of the YRT2 mRNA (as opposed to affecting translation in a general, non-selective or non-specific fashion).

In an example for an assay for an agent that binds to a YRT2 polypeptide, a YRT2 polynucleotide encoding a polypeptide is first recombinantly expressed in a prokaryotic or eukaryotic expression system as a native or as a fusion protein in which the YRT2 polypeptide is conjugated with a well-characterized epitope or protein as described under "Preparation of polypeptides of this invention". Recombinant YRT2 polypeptide is then purified by, for instance, immunoprecipitation using anti-YRT2 polypeptide antibodies or

anti-epitope antibodies or by binding to immobilized ligand of the conjugate. An affinity column made of YRT2 polypeptide or YRT2 polypeptide-fusion protein is then used to screen a mixture of compounds which have been appropriately labeled. Suitable labels include, but are not limited to fluorochromes, radioisotopes, enzymes and chemiluminescent compounds. The unbound and bound compounds can be separated by washes using various conditions (e.g. high salt, detergent) that are routinely employed by those skilled in the art. Non-specific binding to the affinity column can be minimized by pre-clearing the compound mixture using an affinity column containing merely the conjugate or the epitope. Similar methods can be used for screening for an agent(s) that competes for binding to YRT2 polypeptides. In addition to affinity chromatography, there are other techniques such as measuring the change of melting temperature or the fluorescence anisotropy of a protein which will change upon binding another molecule. For example, a BIAcore assay using a sensor chip (supplied by Pharmacia Biosensor, Stitt et al. (1995) *Cell* 80: 661-670) that is covalently coupled to native YRT2 polypeptide, fragments thereof, or YRT2 polypeptide-fusion proteins, may be performed to determine the YRT2 polypeptide binding activity of different agents.

It is understood that anti-YRT2 polypeptide antibodies, including anti-mGluR5 antibodies, can be used to determine whether agents which interact with YRT2 polypeptides alter the structure and/or conformation of the YRT2 polypeptide. For example, a conformational change and/or structural alteration induced by contact with an agent may result in the YRT2 polypeptide becoming unrecognizable by the YRT2 polypeptide-specific antibodies. The loss of the ability of a monoclonal anti-YRT2 polypeptide antibody to immunoprecipitate YRT2 polypeptide after the polypeptide has been contacted by the agent would suggest that the agent had interfered, either directly or through a conformational change and/or a structural alteration, with the antibody recognition site on the YRT2 polypeptide. Other ways to assess this interaction are well known in the art. As such a change may alter YRT2 polypeptide function, agents screened for their effect on YRT2 polypeptide - anti-YRT2 polypeptide interactions would be useful for the refinement of those agents known to interact with YRT2 polypeptide to those that may alter YRT2 polypeptide polynucleotide(s) function. Methods for making antibodies are known in the art and need not be described in detail herein.

Cell-based embodiments

These embodiments employ cell systems containing YRT2 and/or mGluR polynucleotides and/or polypeptide(s) (which, as the definitions make clear, include sequences corresponding to YRT2 polynucleotides and/or polypeptides). In one embodiment, the method provides methods for identifying an agent which may exhibit CNS depressant activity, comprising the steps of (a) introducing a polynucleotide comprising a mouse associated with CNS depressant sensitivity into a suitable host cell, wherein the polynucleotide comprises a mouse polynucleotide sequence corresponding to a polynucleotide sequence contained in yeast artificial chromosome YRT2; (b) contacting host cell of step (a) with at least one agent to be tested; (c) analyze at least one characteristic associated with expression of the polynucleotide, wherein an agent is identified by its ability to modulate expression of the mouse YRT2 polynucleotide.

In another embodiment, the invention provides methods for classifying a CNS depressant comprising the steps of (a) introducing a polynucleotide associated with CNS depressant sensitivity into a suitable host cell, wherein the polynucleotide comprises a mouse polynucleotide sequence corresponding to a polynucleotide sequence contained in yeast artificial chromosome YRT2; (b) contacting the host cell with the CNS depressant to be classified; (c) analyzing at least one characteristic associated with expression of the polynucleotide, wherein a CNS depressant which modulates expression may fall within that class of CNS depressants which displays differential reactivity with respect to LS and SS mice.

These methods could also be used to identify an agent which may modulate CNS depressant activity. In another embodiment, the invention provides methods for identifying an agent which may modulate CNS depressant sensitivity in an individual, comprising (a) introducing a polynucleotide associated with CNS depressant sensitivity into a suitable host cell, wherein the polynucleotide comprises a mouse polynucleotide sequence corresponding to a polynucleotide sequence contained in yeast artificial chromosome YRT2, (b) contacting host cell of step (a) with at least one agent to be tested; (c) analyze at least one characteristic associated with expression of the polynucleotide, wherein an agent is identified by its ability to modulate expression of the mouse YRT2 polynucleotide. Preferably, these methods would compare the effect of an agent alone versus the agent in the presence of a CNS depressant (i.e., contacting the cell with the CNS depressant in addition to the agent).

All of these methods preferably include a control sample which does not receive the agent(s). The characteristic(s) to be analyzed may be assessed in any number of ways, including, but not limited to, microscopic analysis, viability testing, ability to replicate, histological examination, the level of a particular RNA or polypeptide associated with the cells, the level of enzymatic activity expressed by the cells or cell lysates, and the ability of the cells to interact with other cells or compounds. Differences between treated and untreated cells indicate effects attributable to the agent(s). Optimally, the agent has a greater effect on experimental cells than on control cells.

To identify agents that specifically modulate YRT2 polynucleotide transcription, for example, the transcription regulatory regions of a gene contained in YRT2 could be linked to a reporter gene and the construct added to an appropriate host cell. As used herein, the term "reporter gene" means a gene that encodes a gene product that can be identified (i.e., a reporter protein). Reporter genes include, but are not limited to, alkaline phosphatase, chloramphenicol acetyltransferase, β-galactosidase, luciferase and green fluorescence protein (GFP). Identification methods for the products of reporter genes include, but are not limited to, enzymatic assays and fluorimetric assays. Reporter genes and assays to detect their products are well known in the art and are described, for example in Ausubel et al. (1987) and periodic updates. Reporter genes, reporter gene assays and reagent kits are also readily available from commercial sources.

Cells transcribing YRT2 mRNA or polynucleotide could be used to identify agents that specifically modulate the half-life of YRT2 mRNA (or polynucleotide) and/or the translation of YRT2 mRNA. Such cells would also be used to assess the effect of an agent on the processing and/or post-translational modification of a YRT2 polypeptide. An agent could modulate the amount of YRT2 polypeptide in a cell by modifying the turn-over (i.e., increase or decrease the half-life) of the YRT2 polypeptide(s). The specificity of the agent with regard to the YRT2 mRNA and polypeptide would be determined by examining the YRT2 products in the absence of the agent and by examining the products of unrelated mRNAs and polypeptides. Methods to examine mRNA half-life, protein processing, and protein turn-over are well know to those skilled in the art.

Cell-based screening methods could also be useful in the identification of agents that modulate YRT2 polypeptide function through the interaction with a YRT2 polypeptide directly. Such agents could block normal YRT2 polypeptide-ligand interactions, if any, or

could enhance or stabilize such interactions. The effect of the agent could be determined using immunoprecipitation reactions. Anti-YRT2 polypeptide antibodies would be used to precipitate YRT2 polypeptide and any protein tightly associated with it. By comparing the polypeptides immunoprecipitated from treated cells and from untreated cells, an agent could be identified that would augment or inhibit YRT2 polypeptide-ligand interactions, if any. YRT2 polypeptide-ligand interactions could also be assessed using cross-linking reagents that convert a close, but noncovalent interaction between polypeptides into a covalent interaction. Techniques to examine protein-protein interactions are well known to those skilled in the art.

As discussed above, it is understood that these screening methods also encompass using mGluR sequences, including mammalian mGluR5 from mouse, rat, or human, such as human mGluR5.

For methods using mGluR5 sequences, modulation of mGluR5 activity is measured. Accordingly, an agent may modulate mGluR5 function by affecting any of the following non-limiting examples: transcription; translation; post-translational modification; conformation, including conformation of intact receptor; placement in membrane; cellular localization; interaction with other mGluRs; interaction with ligand; interaction with other moieties; altering a function of a member of a signal transduction pathway.

Any of these levels of mGluR5 function may be measured using methods known in the art. Assays for measuring alterations in transcription, translation, and binding have been discussed above. As further example, modulation of mGluR5 activity may be ascertained by measuring the effects on intracellular calcium, inositol phosphate, and/or cyclic AMP (cAMP).

Intracellular calcium concentration can be measured using methods known in the art, such as fura-2. Briefly, recombinant cells expressing mGluR5 (or a functional fragment thereof) are loaded with fura-2 and suspended in buffer containing 0.5 mM $CaCl_2$. An agent(s) is added, and changes in fluorescence signal are measured. Agent(s) may be added over a range of concentrations.

Ion-exchange columns eluted with chloride provide a relatively rapid means of screening for inositol phosphate formation. Inositol phosphate may further be studied using HPLC.

cAMP levels may be measured by heating a sample treated with agent in water at about 70°C for about 5-10 minutes. Cellular debris is removed by centrifugation after cycles of freezing and thawing. cAMP concentration is determined by RIA.

Suitable host cells for these methods include any host cell which is capable of accepting and expressing a YRT2 (or mGluR) polynucleotide. Examples of suitable host cells include eukaryotic cells, such as yeast cells (such as, for example *P. pastoris*, *Saccharomyces cerevisiae, Candidatropicalis, Hansenula polymorpha, Schizosaccharomyces pombe* and the like), insect cells, fungal cells, amphibian cells (such as Xenopus), nemotode cells (such as *C. elegans*) and mammalian cells. Mammalian cells are widely available and need not be discussed in detail herein. Exemplary cells for suitable for practicing these methods include COS cells, mouse L cells, Chinese hamster ovary (CHO) cells, human embryonic kidney (HEK) cells, African green monkey cells, all of mammalian origin; *Xenopus laevis* oöcytes, of amphibian origin; *Saccharomyces cerevisiae, Pichia pastoris* which are yeast. Exemplary cells for expressing injected RNA transcripts include *Xenopus laevis* oöcytes. Methods for injecting Xenopus oöcytes are well-known in the art. Other, more particular, cell types, well-known to those skilled in the art, include HEK293; Ltk^- cells; CDS-7 cells; CHO cells; DG44 cells (dhfr CHO cells; e.g., Urlaub et al. (1986) *Cell. Molec. Genet.* 12:555); BHK cells, and cells of neuronal origin, such as HCN1A andHCN2 (derived from human CNS). Virtually all of these cell types, as well as other suitable cell types, such as neural cells, are either commercially available or are available through the ATCC.

For methods involving mGluR5 sequences, mammalian expression systems, including commercially available systems and other such systems known to those of skill in the art which express G-proteins (either endogenously or recombinantly), for expression of DNA encoding the human metabotropic glutamate receptor subtypes provided herein, are preferred. Xenopus oocytes are preferred for expression of in vitro mRNA transcripts of DNA encoding those human metabotropic receptor subtypes that are coupled to the PI hydrolysis/Ca^{++} signalling pathways. An endogenous inositol triphosphate second messenger-mediated pathways in oocytes permits functional expression of human metabotropic receptors in these cells. Oocytes expressing recombinant human metabotropic receptors respond to agonist via the oocyte G-protein-coupled IP_3 generation pathway, which stimulates release of Ca^{++} from internal stores, and reportedly activates a

chloride channel that can be detected as a delayed oscillatory current by voltage-clamp recording. Methods for injecting Xenopus oöcytes are well-known in the art.

Host cells for functional recombinant expression of human metabotropic receptors preferably express endogenous or recombinant guanine nucleotide-binding proteins (i.e., G-proteins). G-proteins are a highly conserved family of membrane-associated proteins composed of α, β and γ subunits. The α subunit, which binds GDP and GTP, differs in different G-proteins. The attached pair of β and γ subunits may or may not be unique; different α claims may be linked to an identical αγ pair or to different pairs. Linder and Gilman (1992) *Sci. Am.* 267:56-65. More than 30 different cDNAs encoding G protein α subunits have been cloned. See, for example, Simon et al. (1991) *Science* 252:802. Four different β polypeptide sequences are known. Simon et al. (1991). Three of five identified γcDNAs have been cloned. Hurley et al. (1984) *Proc. Natl. Acad. Sci. U.S.A.* 81:6948; Gautam et al. (1989) *Science* 244:971; and Gautam et al.(1990) *Proc. Natl. Acad. Sci. U.S.A.* 87:7973. The sequences of a fourth γ cDNA (Kleuss et al. (1993) *Science* 259:832) and a fifth γ cDNA (Fisher and Aronson (1992) *Mol. Cell. Bio.* 12:1585) have been established, and additional γ subtypes may exist (Tamir et al. (1991) *Biochemistry* 30:3929. G-proteins switch between active and inactive states by guanine nucleotide exchange and GTP hydrolysis. Inactive G protein is stimulated by a ligand activated receptor to exchange GDP for GTP. In the active form, the α subunit, bound to GTP, dissociates from the βγ complex, and the subunits then interact specifically with cellular effector molecules to evoke a cellular response. Because different G-proteins can interact with different effector systems (e.g., phospholipase C, adenyl cyclase systems) and different receptors, it is useful to investigate different host cells for expression of different recombinant human metabotropic receptor subtypes. Alternatively, host cells can be transfected with G-protein subunit-encoding DNAs for heterologous expression of differing G proteins.

A practitioner of ordinary skill will be well acquainted with techniques for transfecting eukaryotic cells, including the preparation of a suitable vector, such as a viral vector; conveying the vector into the cell, such as by electroporation; and selecting cells that have been transformed, such as by using a reporter or drug sensitivity element. The effect of an agent on transcription from a YRT2 regulatory region in these constructs would be assessed through the activity of the reporter gene product.

DNA may be stably incorporated into cells or may be transiently expressed using methods known in the art. Stably transfected mammalian cells may be prepared by transfecting cells with an expression vector having a selectable marker gene (such as, for example gene for thymidine kinase, dihydrofolate reductase, neomycin resistance, and the like), and growing the transfected cells under conditions selective for cells expressing the marker gene. To prepare transient transfectants, mammalian cells are transfected with a reporter gene (such as the *E. coli β-galactosidase* gene) to monitor transfection efficiency. Selectable marker genes are typically not included in the transient transfections because the transfectants are typically not grown under selective conditions, and are usually analyzed within a few days after transfection.

In vivo embodiments

The invention also includes methods for identifying agents based on transgenic, or in vivo, systems. For these methods, an agent is administered to a transgenic animal, such as a mouse, harboring, for example, mouse sequences corresponding to YRT2 as described above. The agent is assessed by analyzing a characteristic associated with CNS depressant activity, such as, for example, loss of righting reflex (LORR), or sleep times. Various agents may be assessed for their ability or inability to affect any of these characteristics. Various agents may also be assessed for their ability to alter the effect of a CNS depressant, thus indicating that this agent(s) may modulate CNS depressant activity.

Accordingly, the invention includes methods of identifying an agent which may exhibit CNS depressant activity, comprising the steps of (a) adminstering an agent to a transgenic animal containing an introduced mouse YRT2 sequence; and (b) analyzing at least one characteristic associated with CNS depressant activity, wherein an agent is identified by its ability to modulate the characteristic. The invention also includes methods of classifying agents and methods of identifying an agent which may modulate CNS depressant activity using these steps. For any of these transgenic embodiments, characteristics which are associated with CNS depressant activity include, but are not limited to, sleep (unconsciousness) time, degree of loss of reflex or response to stimuli, and degree of loss of pain response. Alternatively, the transgenic embodiments may employ mGluR sequences as described above.

Preferably, these methods are performed by comparing the effect of the agent(s) on the transgenic animal described above with the effect on a SS (resistant) mouse. Candidate

CNS depressant agents may also be thus classified by making this comparison with respect to whether they belong in the class of CNS depressants which display differential activity in LS versus SS mice.

In one embodiment, the transgenic animal (such as mouse) contains mGluR5 sequences. These mGluR5 sequences may be from a mammal, such as mouse, rat, primate or human. Alternatively, the transgenic animal contains other mGluR sequences, such as mGluR1, mGluR2, mGluR3, mGluR4, mGluR6, mGluR7 and mGluR8 (as well as any splice variants or non-mammalian mGluR homologs). Any of these various mGluR sequences may be from mammalian systems, such as mouse, rat, primate or human, or non-mammalian systems, such as homologous sequences from Drosophila and *C. elegans*.

Obtaining these sequences has been described above. A transgenic animal may be mammals, such as mice, rats, primate, cows, pigs, dogs and cats. Alternatively, a transgenic animal may be an invertebrate, such as Drosophila or C. elegans (nemotode). Methods of making transgenic animals, particularly mice, are known in the art. Briefly, polynucleotide sequence(s) are isolated using standard methods of the art. Purified polynucleotide (usually DNA) is injected or introduced into appropriate early cell types or lineages, such as oocytes. See, for example, Brinster et al.(1985) *Proc. Natl. Acad. Sci. USA* 82:4438-4442. Other means known in the art of introducing the polynucleotide are transfection and transduction. Alternatively, appropriate cells may be infected with viruses, such as retroviruses of adenoviruses, which are modifed to carry the polynucleotide sequences to be introduced.

Pluripotent stem cells derived from the inner cell mass of the embryo and stabilized in culture can be manipulated in culture to incorporate nucleotide sequences. A transgenic animal can be produced from such stem cells through implantation into a blastocyst that is implanted into a foster mother an allowed to come to term. In the case of mice, injected oocytes are transferred into pseudopregnant females. Animals suitable for transgenic experiments can be obtained from standard commercial sources such as Charles River (Wilmington, MA) and Harlan Sprague Dawley (Indianapolis, IN).

Methods for culturing embryonic stem (ES) cells and the subsequent production of transgenic animals by the introduction of DNA into ES cells using methods such as electroporation, calcium phosphate/DNA precipitation and direct injection are well known in the art, as are procedure for embryo manipulations. Microinjection procedures for fish,

amphibian eggs and birds are detailed by Houdebine and Chourrout (1991) *Experientia* 47:897-905. See also U.S. Pat. 4,945,050.

Transfection and isolation of desired clones can be carried out using standard techniques. For example, random gene integration can be carried out by co-transfecting the nucleic acid with a gene encoding antibiotic resistance. Alternatively, the gene encoding antibiotic resistance is physically linked to a polynucleotide used in the screening methods described herein.

DNA molecules introduced into ES cells can also be integrated into the chromosome through homologous recombination. See, for example, Capecchi (1989) *Science* 244:1288-1292. Methods for positive selection of the recombination event (such as neomycin resistance) and dual positive-negative selection (e.g., neomycin resistance and gancyclovir resistance) and the subsequent identification of desired clones by PCR are methods standard in the art. Capecchi (1989); Joyner et al. (1989) *Nature* 338:153-156.

Targeted ES cells are then injected into blastocysts, which are transferred into pseudopregnant females. Resulting chimeric animals are bred and offspring analyzed by Southern blotting to identify individuals carrying the introduced polynucleotide sequences.

Transgenic mice are identified by preparing DNA from, for example, tail biopsies and simultaneous hybridization with appropriate probes that indicate the presence of introduced DNA. In certain circumstances, presence of introduced DNA (polynucleotide) is indicated by a measurable phenotype, such as change in pigmentation. The transgenic animal (herein exemplified by mouse, due to its more common use) may or may not have more than one copy of the YRT2 sequence(s) or mGluR sequence(s). Depending on the nature of the experiment, it may also be desirable to generate different transgenic lines which have different copy numbers. Copy number can be determined by using methods standard in the art, such as Southern blot analysis or PCR.

The procedure for generating transgenic rats is similar to that of mice. See, for example, Hammer et al. (1990) *Cell* 63:1099-1112. Procedures for production of transgenic non-rodent mammals and other animals are also known in the art. See, for example, Houdebine and Chourrout; Pursel et al. (1989) *Science* 244:1281-1288; Simms et al. (1988) *Bio/Technology* 6:179-183.

If desired, the endogenous mGluR gene may be inactivated. This may be accomplished, for example, by using a recombinant gene engineered to contain an

insertional mutation (such as *neo*). The recombinant gene is inserted into the genome of a recipient cells, tissue, or animal, and transcription of the endogenous mGluR is prevented, or decreased.

Polynucleotide and polypeptide sequences of the invention

The invention provides new polynucleotide and polypeptide sequences associated with differential CNS depressant response, namely, a class of CNS depressants defined by a differential response/sensitivity in LS versus SS mice. Examples of these CNS depressants are discussed above. In particular, the invention provides mouse mGluR5 polynucleotide and polypeptide sequences. These sequences have a variety of uses, including use in screening methods described above, as indicators of CNS depressant function, and use in obtaining corresponding human sequences which are associated with differential response to CNS depressants, which in turn may provide a useful drug target.

The polynucleotide sequence of most of mGluR5 coding region is depicted in SEQ ID NO:1. The polypeptide sequence of most of mGluR5 is depicted in SEQ ID NO:2.

Bacteria containing respective cloned DNA have been deposited with the American Type Culture Collection (ATCC), 12310 Parklawn Drive, Rockville, Maryland, U.S.A. 20852 on _____, 199_, under the provisions of the Budapest Treaty for the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. They were accorded Accession Numbers as follows:

_____.

Polynucleotides of the invention

The invention provides mGluR5 polynucleotides, which will be described below. The invention also provides vectors containing these polynucleotides, host cells containing these polynucleotides, and compositions comprising these polynucleotides. These polynucleotides are isolated and/or produced by chemical and/or recombinant methods, or a combination of these methods. Unless specifically stated otherwise, "polynucleotides" shall include all embodiments of the polynucleotides of this invention. It is also understood that, all polynucleotide embodiments are isolated polynucleotides. The cloning of mGluR5 polynucleotide sequences is described in Example 4. Examination of Genbank using a BLAST search revealed that a portion of mouse mGluR5, represented by an exon, had previously been obtained (SEQ ID NO:3). However, the conceptual translation of this

region does not correspond to the conceptual amino acid translation of SEQ ID NO:1, which is depicted in SEQ ID NO:2.

Accordingly, this invention provides an isolated polynucleotide that contains a sequence encoding a mouse mGluR5 polypeptide wherein the polypeptide is at least about 10 amino acids in length and is depicted in SEQ ID NO:2 but not depicted in SEQ ID NO:6 (rat mGluR5 amino acid sequence) or SEQ ID NO:7 (human mGluR5 amino acid sequence). In other embodiments, the mouse mGluR5 encoded is at least about 20, at least about 25, at least about 50, at least about 75, at least about 100, at least about 125, at least about 200, at least about 250, at least about 300, at least about 350, at least about 400, at least about 500, at least about 700, at least about 750, at least about 1000 amino acids in length and depicted in SEQ ID NO:2 but not depicted in SEQ ID NO:6 or SEQ ID NO:7. The invention also includes an isolated polynucleotide comprising a polynucleotide encoding the polypeptide of SEQ ID NO:2.

In other embodiments, the invention includes an isolated polynucleotide comprising a polynucleotide of at least about 10 contiguous nucleotides of SEQ ID NO:1, wherein the about 10 contiguous nucleotides are not depicted in SEQ ID NO:3 (the above-discussed mouse mGluR5 sequence on the Genbank database) or SEQ ID NO:5 (rat mGluR5 polynucleotide sequence). Alternatively, the invention includes an isolated polynucleotide comprising a polynucleotide of at least about 10 contiguous nucleotides of nucleotide 1 to 2625 SEQ ID NO:1, wherein the 10 contiguous nucleotides are not depicted in SEQ ID NO:5. The invention also includes an isolated polynucleotide comprising a polynucleotide of at least about 10 contiguous nucleotides of nucleotide 2726 to 3591 of SEQ ID NO:1, wherein the 10 contiguous nucleotides are not depicted in SEQ ID NO:5. Alternatively, the isolated polynucleotide comprises a polynucleotide of at least about 25, at least about 50, at least about 75, at least about 100, at least about 150, at least about 200, at least about 250, at least about 300, at least about 500, at least about 750, at least about 1000 contiguous nucleotides of SEQ ID NO:1, wherein the contiguous nucleotides are not depicted in SEQ ID NO:5. In another embodiment, the invention includes a polynucleotide comprising the sequence of SEQ ID NO:1. The shorter embodiments are useful, for example, as probes as well as providing a template for production of mGluR5 polypeptide(s) and/or functional fragments.

The invention includes modifications to the mGluR5 polynucleotides described above such as deletions, substitutions, additions, or changes in the nature of any nucleic acid moieties. A "modification" is any difference in nucleotide sequence as compared to a polynucleotide shown herein to encode an mGluR5 polypeptide, and/or any difference in terms of the nucleic acid moieties of the polynucleotide(s). Such changes can be useful to facilitate cloning and modifying expression of mGluR5 polynucleotides. Such changes also can be useful for conferring desirable properties to the polynucleotide(s), such as stability. The definition of polynucleotide provided herein gives examples of these modifications. Hence, the invention also includes functionally-preserved variants of the nucleic acid sequences disclosed herein, which include nucleic acid substitutions, additions, and/or deletions.

The invention also encompasses mGluR5 polynucleotides including full-length (unprocessed), processed, coding, non-coding (including flanking region) or portions thereof, provided that these polynucleotides contain a region encoding at least a portion of mGluR5. Also embodied are the mRNA and cDNA sequences and fragments thereof that include a portion mGluR5 encoding segment.

The invention also encompasses polynucleotides encoding for functionally equivalent variants and derivatives of full-length mGluR5 and functionally equivalent fragments thereof which may enhance, decrease or not significantly affect properties of the polypeptides encoded thereby. For instance, changes in a DNA sequence that do not change the encoded amino acid sequence, as well as those that result in conservative substitutions of amino acid residues, non-deleterious non-conservative substitutions, one or a few amino acid deletions or additions, and substitution of amino acid residues by amino acid analogs are those which will not significantly affect properties of the encoded polypeptide. Nucleotide substitutions that do not alter the amino acid residues encoded can be useful for optimizing gene expression in different systems. Suitable substitutions are known to those of skill in the art and are made, for instance, to reflect preferred codon usage in the particular expression systems. In another example, alternatively spliced polynucleotides can give rise to a functionally equivalent fragment or variant of mGluR5. Alternatively processed polynucleotide sequence variants are defined as polynucleotide sequences corresponding to mRNAs that differ in sequence for one another but are derived from the same genomic region, for example, mRNAs that result from: 1) the use of

alternative promoters; 2) the use of alternative polyadenylation sites; and/or 3) the use of alternative splice sites.

The mGluR5 polynucleotides of the invention also include polynucleotides encoding other mGluR5 fragments. The polynucleotides encoding mGluR5 fragments are useful, for example, as probes, therapeutic agents, a polypeptide processing signal, and as a template for encoding various functional domains of mGluR5. Accordingly, the invention includes a polynucleotide that comprises a region of at least 15 contiguous nucleotides, more preferably at least about 20 contiguous nucleotides, more preferably at least about 25 contiguous nucleotides, more preferably at least about 35 contiguous nucleotides, more preferably at least about 50 contiguous nucleotides, even more preferably at least about 75 contiguous nucleotides, even more preferably at least about 100 contiguous nucleotides, even more preferably at least about 200 contiguous nucleotides, even more preferably at least about 300 contiguous nucleotides, even more preferably at least about 500 contiguous nucleotides, even more preferably at least about 750 contiguous nucleotides, even more preferably at least about 1000 contiguous nucleotides, even more preferably at least about 1250 contiguous nucleotides, of SEQ ID NO:1, provided that the contiguous nucleotides are not depicted in SEQ ID NO:3 or SEQ ID NO:5.

Another embodiment of the invention is isolated polynucleotides comprising a region of at least about 20 contiguous nucleotides, with the region having at least 97% sequence identity with a sequence depicted in SEQ ID NO:1. Alternatively, the region may also have 98% sequence identity, preferably 98.5% sequence identity, preferably 99% sequence identity. Alternatively, these regions may comprise at least about 25 contiguous nucleotides, at least about 30 contiguous nucleotides, at least about 50 contiguous nucleotides, at least about 60 contiguous nucleotides, at least about 75 contiguous nucleotides, at least about 100 contiguous nucleotides, at least about 200 contiguous nucleotides, at least about 250 contiguous nucleotides, at least about 300 contiguous nucleotides, at least about 500 contiguous nucleotides, at least about 700 contiguous nucleotides, at least about 1000 contiguous nucleotides, any and each of which can have any of the above-specified degrees of sequence identity. It is understood that these contiguous nucleotide sequences are not depicted in SEQ ID NO:3 or SEQ ID NO:5.

As known to those of skill in the art, methods for determination of sequence identity between two polynucleotides can include the use of hybridization reactions and sequence

alignment algorithms. Alignment of known polynucleotide sequences relative to each other with commercially available sequence comparison programs, such as BLAST programs or those found in the GCG, Inc. software packages, provide a means to directly compare sequences for degree of identity. Such programs establish the sequence alignments and can provide the percentage of sequence identity between them.

Generally, such programs identify the best sequence alignment by scoring the alignment for matches, mismatches, and gaps in the sequence created in the generation of the alignment. For example, the BESTFIT program from GCG uses the algorithm of Smith and Waterman with the default scoring of a match value of 10, a mismatch value of -9, a gap creation penalty of 50 and a gap extension penalty of 3. Thus, the quality score for a polynucleotide alignment can be determined using the equation:

Quality = (10 x matches) + (-9 x mismatches) - (50 x gap number) - (3 x total length of gaps)

Other alignment programs are based on other algorithms which score alignment with different values (e.g., the GAP program from GCG uses the Needleman and Wunsch algorithm with the default match and gap scoring values as desribed for the BESTFIT program but a default mismatch value of 0 and the FASTA program uses a Pearson and Lipman search with a default gap creation penalty of 16 and gap extension penalty of 4).

Similar alignment programs can be used to assess amino acid sequence similarity between polypeptide sequences. Often, the alignment of polypeptide sequences are scored using values different than those used with polynucleotides. For example, the FASTA program uses a default gap creation penalty of 12 and gap extension penalty of 2. Once the polypeptides are aligned relative to each other, the programs can provide the degree of amino acid similarity and identity.

Hybridization can also be used in the determination of polynucleotide sequence identity. The stability of hybridized sequences is reflected in the melting temperature (T_m; discussed below) of the hybrids. For polynucleotides greater than 100 nucleotides in length, the T_m can be used to approximate the degree of sequence identity between two sequences because the T_m appears to decrease 0.5 - 1.5 °C for every 1% base pair mismatch in the hybrid. This approach can only approximate the degree of identity because the base composition of and mismatch distribution in the hybrids influence the actual change in T_m.

Generally, to determine the approximate sequence identity between two polynucleotide sequences with hybridization, the lowest temperature at which hybridization occurs between identical sequences is established for a particular hybridization solution (see below) and the T_m is determined. When the polynucleotide sequence is then hybridized with another sequence, a decrease in the T_m of the reaction can be correlated to the degree of mismatch between the sequences as described above (i.e., roughly 1% degree of mismatch reduces the T_m of a heteroduplex by 0.5 - 1.5 °C).

Hybridization reactions can be performed under conditions of different "stringency". Stringency of hybridization is used herein to refer to conditions under which polynucleotide hybrids are stable. Conditions that increase stringency of a hybridization reaction of widely known and published in the art. See, for example, Sambrook et al. (1989) and Ausubel et al. (1987).

"T_m" is the temperature in degrees Centigrade at which 50% of a polynucleotide duplex made of complementary strands hydrogen bonded in anti-parallel direction by Watson-Crick base pairing dissociates into single strands under conditions of the experiment. T_m may be predicted according to a standard formula, such as:

$$T_m = 81.5 + 16.6 \log[X^+] + 0.41 (\%G/C) - 0.61 (\%F) - 600/L$$

where $[X^+]$ is the cation concentration (usually sodium ion, Na^+) in mol/L; (%G/C) is the number of G and C residues as a percentage of total residues in the duplex; (%F) is the percent formamide in solution (wt/vol); and L is the number of nucleotides in each strand of the duplex. Thus, hybrid stability is dependent upon a number of factors including, but not limited to, ionic strength of the hybridization and wash solutions, base composition of the polynucleotides involved in the duplex; destabilizing agents in the hybridization solution (e.g., formamide or urea),and length of the duplex formed.

In terms of hybridization conditions, the higher the sequence identity required, the more stringent are the hybridization conditions if such sequences are determined by their ability to hybridize to a sequence of SEQ ID NO:1. Accordingly, the invention also includes polynucleotides that are able to hybridize to a sequence comprising at least 20 contiguous nucleotides (or more, such as 25, 35, 50, 75 or 100 contiguous nucleotides) of SEQ ID NO:1. The hybridization conditions would be stringent, e.g., 80°C (or higher temperature) and 6 X SSC (or less concentrated SSC).

Examples of relevant conditions include (in order of increasing stringency): incubation temperatures of 25°C, 37°C, 50°C and 68°C; buffer concentrations of 10 X SSC, 6 X SSC, 1 X SSC, 0.1 X SSC (where 1 X SSC is 0.15 M NaCl and 15 mM citrate buffer) and their equivalents using other buffer systems; formamide concentrations of 0%, 25%, 50%, and 75%; incubation times from 24 hours to 5 minutes; 1, 2, or more washing steps; wash incubation times of 1, 2, or 15 minutes; and wash solutions of 6 X SSC, 1 X SSC, 0.1 X SSC, or deionized water, in temperatures of 25°C, 37°C, 50°C and 68°C.

Accordingly, the invention includes an isolated polynucleotide capable of hybridizing under moderately stringent conditions, more preferably under stringent conditions, to a region of SEQ ID NO:1, wherein the region is at least about 10 contiguous nucleotides. The region may also be at least about 25 contiguous nucleotides, at least about 50 contiguous nucleotides, at least about 100 contiguous nucleotides, at least about 150 contiguous nucleotides, at least about 200 contiguous nucleotides, at least about 200 contiguous nucleotides, at least about 300 contiguous nucleotides, at least about 500 contiguous nucleotides, at least about 1000 contiguous nucleotides of a region of SEQ ID NO:1. The invention also includes an isolated polynucleotide capable of hybridizing under moderately stringent conditions, even more preferably under stringent conditions, to the sequence depicted in SEQ ID NO:1.

Compositions containing mGluR5 polynucleotides are encompassed by this invention. The invention also provides compositions comprising a vector(s) containing an mGluR5 polynucleotide as well as compositions comprising a host cell containing an mGluR5 polynucleotide, as described herein. Generally, the compositions further contain a physiologically acceptable medium, such as water, buffered solutions, or pharmaceutically acceptable excipient. These media are well-known in the art.

Preparation of mGluR5 *polynucleotides of this invention*

The polynucleotides of this invention can be obtained using chemical synthesis, recombinant methods, or PCR.

Methods of chemical polynucleotide synthesis are well known in the art and need not be described in detail herein. One of skill in the art can use the sequences provided herein and a commercial DNA synthesizer to produce a desired DNA sequence.

For preparing mGluR5 polynucleotides using recombinant methods, a polynucleotide comprising a desired sequence can be inserted into a suitable vector, and the

vector in turn can be introduced into a suitable host cell for replication and amplification. Polynucleotides may be inserted into host cells by any means known in the art. Cells are transformed by introducing an exogenous polynucleotide by direct uptake, endocytosis, transfection, F-mating or electroporation. Once introduced, the exogenous polynucleotide can be maintained within the cell as a non-integrated vector (such as a plasmid) or integrated into the host cell genome. The polynucleotide so amplified can be isolated from the host cell by methods well known within the art. See, e.g., Sambrook et al. (1989).

Alternatively, PCR allows reproduction of DNA sequences. PCR technology is well known in the art and is described in U.S. Pat. Nos. 4,683,195, 4,800,159, 4,754,065 and 4,683,202, as well as *PCR: The Polymerase Chain Reaction*, Mullis et al. eds., Birkauswer Press, Boston (1994).

RNA can be obtained by using the isolated DNA in an appropriate vector and inserting it into a suitable host cell. When the cell replicates and the DNA is transcribed into RNA, the RNA can then be isolated using methods well known to those of skill in the art, as set forth in Sambrook et al., (1989), for example. RNA can also be obtained through in vitro reactions. For example, the mGluR5 polynucleotide can be inserted into a vector that contains appropriate transcription promoter sequences. Commercially available RNA polymerases will specifically initiate transcription at their promoter sites and continue the transcription process through the adjoining DNA polynucleotides. Placing the mGluR5 polynucleotides between two such promoters allows the generation of sense or antisense strands of mGluR5 RNA.

If used as a vaccine (i.e., pharmaceutical composition for eliciting an immune response), plasmids containing mGluR5 polynucleotides are preferably prepared as described by Horn et al. ((1995) *Human Gene Therapy* 6:565-573) which produces a pharmaceutical grade plasmid DNA suitable for administration.

Cloning and expression vectors comprising mGluR5 polynucleotide

The present invention further includes a variety of vectors containing mGluR5 polynucleotides of this invention. These vectors can be used for expression of recombinant polypeptides as well as a source of mGluR5 polynucleotides. Cloning vectors can be used to obtain replicate copies of the mGluR5 polynucleotides they contain, or as a means of storing the polynucleotides in a depository for future recovery. Expression vectors (and host cells containing these expression vectors) can be used to obtain polypeptides produced

from the polynucleotides they contain. They may also be used where it is desirable to express mGluR5 polypeptides in an individual, such as for eliciting an immune response via the polypeptide(s) encoded in the expression vector(s). Suitable cloning and expression vectors include any known in the art, e.g., those for use in in vitro, bacterial, mammalian, yeast and insect expression systems. Specific vectors and suitable host cells are known in the art and need not be described in detail herein. For example, see Gacesa and Ramji, *Vectors*, John Wiley & Sons (1994).

Cloning and expression vectors typically contain a selectable marker (for example, a gene encoding a protein necessary for the survival or growth of a host cell transformed with the vector), although such a marker gene can be carried on another polynucleotide sequence co-introduced into the host cell. Only those host cells into which a selectable gene has been introduced will survive and/or grow under selective conditions. Typical selection genes encode protein(s) that (a) confer resistance to antibiotics or other toxins substances, e.g., ampicillin, neomycyin, methotrexate, etc.; (b) complement auxotrophic deficiencies; or (c) supply critical nutrients not available from complex media. The choice of the proper marker gene will depend on the host cell, and appropriate genes for different hosts are known in the art. Cloning and expression vectors also typically contain a replication system recognized by the host.

Suitable cloning vectors may be constructed according to standard techniques, or may be selected from a large number of cloning vectors available in the art. While the cloning vector selected may vary according to the host cell intended to be used, useful cloning vectors will generally have the ability to self-replicate, may possess a single target for a particular restriction endonuclease, and/or may carry genes for a marker that can be used in selecting clones containing the vector. Suitable examples include plasmids and bacterial viruses, e.g., pUC18, mp18, mp19, pBR322, pMB9, ColE1, pCR1, RP4, phage DNAs, and shuttle vectors such as pSA3 and pAT28. These and many other cloning vectors are available from commercial vendors such as BioRad, Stratagene, and Invitrogen.

Expression vectors generally are replicable polynucleotide constructs that contain a polynucleotide encoding an mGluR5 polypeptide of interest. The mGluR5 polynucleotide encoding the polypeptide is operatively linked to suitable transcriptional controlling elements, such as promoters, enhancers and terminators. For expression (i.e., translation), one or more translational controlling elements are also usually required, such as ribosome

binding sites, translation initiation sites, and stop codons. These controlling elements (transcriptional and translational) may be derived from the mGluR5 gene, or they may be heterologous (i.e., derived from other genes and/or other organisms). A polynucleotide sequence encoding a signal peptide can also be included to allow an mGluR5 polypeptide to cross and/or lodge in cell membranes or be secreted from the cell. A number of expression vectors suitable for expression in eukaryotic cells including yeast, avian, and mammalian cells are known in the art. Another example of an expression vector (system) is the baculovirus/insect cell system. Expression of mGluR5 RNA in vitro is described above.

The vectors containing the polynucleotides of interest can be introduced into the host cell by any of a number of appropriate means, including electroporation, transfection employing calcium chloride, rubidium chloride, calcium phosphate, DEAE-dextran, or other substances; microprojectile bombardment; lipofection; and infection (where the vector is an infectious agent, such as vaccinia virus). The choice of means of introducing vectors or mGluR5 polynucleotides will often depend on the host cell.

Host cells transformed with mGluR5 polynucleotides

Another embodiment of this invention are host cells transformed with (i.e., comprising) mGluR5 polynucleotides and/or vectors having mGluR5 polynucleotide(s) sequences, as described above. Both prokaryotic and eukaryotic host cells may be used. Prokaryotic hosts include bacterial cells, for example *E. coli*, *B. subtilis*, and mycobacteria. Among eukaryotic hosts are yeast, insect, avian, plant, C. elegans (or nematode) and mammalian cells. Host systems are known in the art and need not be described in detail herein. Examples of fungi (including yeast) host cells are *S. cerevisiae, Kluyveromyces lactis (K. lactis),* species of *Candida* including *C. albicans* and *C. glabrata, Aspergillus nidulans, Schizosaccharomyces pombe (S. pombe), Pichia pastoris*, and *Yarowia lipolytica.* Examples of mammalian cells are COS cells, house L cells, Chinese hamster ovary (CHO) cells, human embroyonic kidney (HEK) cells, African green monkey cells. Xenopus laevis oocytes, or other cells of amphibian origin, may also be used.

The host cells of this invention can be used, *inter alia*, as repositories of mGluR5 polynucleotides and/or vehicles for production of mGluR5 polynucleotides and/or polypeptides. They may also be used in the screening methods described above.

Polypeptides of the invention

The present invention encompasses mouse mGluR5 polypeptide sequences. Unless specifically stated, the term "polypeptide(s)" shall include all polypeptide embodiments of this invention. It is also further understood that polypeptide embodiments of this invention are isolated polypeptides.

The polypeptides have a variety of uses, including their use in making antibodies that bind to these polypeptides, their use as agents to screen pharmaceutical candidates (both in vitro and in vivo), and their use in rational (i.e., structure-based) drug design. The mGluR5 polypeptides may also be used to identify proteins that interact physically with mGluR5 which could thus themselves be drug targets.

The amino acid sequence of mGluR5 is shown in SEQ ID NO:2. The publicly available conceptual translation of SEQ ID NO:3 (depicted in SEQ ID NO:4) does not correspond to any region of the conceptual translation of SEQ ID NO:1 (as depicted in SEQ ID NO:2).

In one embodiment, the invention includes an isolated polypeptide comprising at least about 5 contiguous amino acids of the sequence of SEQ ID NO:2, wherein the about 5 contiguous amino acids are not depicted in SEQ ID NO:6 or in SEQ ID NO:7, and wherein the isolated polypeptide exhibits an mGluR5 function. In this context, "mGluR5 function" includes, not only those functions delineated in the "Definitions" section of this application, but other functions such as ability to elicit an immune response, including a humoral and/or cellular immune response. In other embodiments, the isolated polypeptide comprises at least about 10 contiguous amino acids, at least about 15 contiguous amino acids, at least about 20 contiguous amino acids, at least about 25 contiguous amino acids, at least about 30 contiguous amino acids, at least about 50 contiguous amino acids, at least about 75 contiguous amino acids, at least about 100 contiguous amino acids, at least about 150 contiguous amino acids, at least about 200 contiguous aminoacids, at least about 250 contiguous amino acids, at least about 300 contiguous amino acids, at least about 400 contiguous amino acids, at least about 500 contigous amino acids of the sequence of SEQ ID NO:2. In any or all of these instances, the contiguous amino acids are not depicted in SEQ ID NO:6 or SEQ ID NO:7.

In another embodiment, the invention includes an isolated polypeptide comprising an mGluR5 polypeptide , wherein the polypeptide comprises the sequence of SEQ ID NO:2.

In another embodiment, the invention provides an isolated polypeptide comprising at least 10 contiguous amino acids which have at least 97% sequence identity to a sequence depicted in SEQ ID NO:2. In other embodiments, the isolated polypeptide comprises at least at least about 15 contiguous amino acids, at least about 20 contiguous amino acids, at least about 25 contiguous amino acids, at least about 30 contiguous amino acids, at least about 50 contiguous amino acids, at least about 75 contiguous amino acids, at least about 100 contiguous amino acids, at least about 150 contiguous amino acids at least about 200 contiguous aminoacids, at least about 250 contiguous amino acids, at least about 300 contiguous amino acids, at least about 400 contiguous amino acids, at least about 500 contigous amino acids, any of which having at least a 98%, preferably at least about 98.5%, preferably at least about 99% sequence identity to a sequence of SEQ ID NO:2. In any or all of these instances, the contiguous amino acids are not depicted in SEQ ID NO:6 or SEQ ID NO:7.

Determination of sequence identity in the context of polypeptide sequences generally involves alignment programs readily and commercially available, and have been discussed above.

The size of mGluR5 polypeptides may vary widely, as the length required to effect activity could be as small as, for example, a 5- or 10-mer amino acid sequence to effect binding and trigger a response. The maximum length generally is not detrimental to effecting activity. The minimum size must be sufficient to provide a desired function. Thus, the invention includes polypeptide fragments of mGluR5 comprising a portion of the amino acid sequence depicted in SEQ ID NO:2 in which the mGluR5 polypeptide is about 15, preferably 25, more preferably 50 more preferably 75, more preferably 100 amino acids in length. These lengths could depend on the proposed use of the mGluR5 polypeptide; for example, a polypeptide comprising the transmembrane domain of mGluR5 could be desired, or the extracellular and/or intracellular domain. Based on the amino acid sequence, and what is known about domains of other mGluRs, these domains could be estimated. Alternatively, it may be that only a partial extracellular domain is necessary if binding of glutamate is the

desired function. As noted above, only a relatively small polypeptide could be used if binding to antibody or eliciting an immune response is desired.

As is evident to one skilled in the art, these mGluR5 polypeptides, regardless of their size, may also be associated with, or conjugated with, other substances or agents to facilitate, enhance, or modulate function and/or specificity of an mGluR5 polypeptide.

The invention includes modifications to mGluR5 polypeptides including functionally equivalent fragments of the mGluR5 polypeptides which do not significantly affect their properties and variants which have enhanced or decreased activity. Collectively, these modifications may be termed "analogs" of mGluR5 or a fragment of mGluR5. Modification of polypeptides is routine practice in the art and need not be described in detail herein. Examples of modified polypeptides include polypeptides with conservative substitutions of amino acid residues, one or more deletions or additions of amino acids which do not significantly deleteriously change the functional activity, or use of chemical analogs. Amino acid residues which can be conservatively substituted for one another include but are not limited to: glycine/alanine; valine/isoleucine/leucine; asparagine/glutamine; aspartic acid/glutamic acid; serine/threonine; lysine/arginine; and phenylalanine/tryosine. Such conservative substitutions are known in the art, and preferably, the amino acid substitutions would be such that the substituted amino acid would possess similar chemical properties as that of the original amino acid. These polypeptides also include glycosylated and non-glycosylated polypeptides, as well as polypeptides with other post-translational modifications, such as, for example, glycosylation with different sugars, acetylation, and phosphorylation. Amino acid modifications can range from changing or modifying one or more amino acids to complete redesign of a region. Other methods of modification include using coupling techniques known in the art, including, but not limited to, enzymatic means, oxidative substitution and chelation. Modifications can be used, for example, for attachment of labels for immunoassay, such as the attachment of radioactive moieties for radioimmunoassay. Modified mGluR5 polypeptides are made using established procedures in the art and can be screened using standard assays known in the art.

The invention also encompasses fusion proteins comprising one or more mGluR5 polypeptides. For purposes of this invention, an mGluR5 fusion protein contains one or more mGluR5 polypeptides and another amino acid sequence to which it is not attached in

the native molecule, for example, a heterologous sequence or a homologous sequence from another region. Useful heterologous sequences include, but are not limited to, sequences that provide for secretion from a host cell, enhance immunological reactivity, or facilitate the coupling of the polypeptide to an immunoassay support or a vaccine carrier. For instance, an mGluR5 polypeptide can be fused with a bioresponse modifier. Examples of bioresponse modifiers include, but are not limited to, cytokines or lymphokines such as GM-CSF, interleukin-2 (IL-2), interleukin 4 (IL-4), and γ-interferon. Accordingly, the invention includes mGluR5 fusion polypeptides that contain GM-CSF or IL-2. Another useful heterologous sequence is one which facilitates purification. Examples of such sequences are known in the art and include those encoding epitopes such as Myc, HA (derived from influenza virus hemagglutinin), His-6, or FLAG. Other heterologous sequences that facilitate purification are derived from proteins such as glutathione S-transferase (GST), maltose-binding protein (MBP), or the Fc portion of immunoglobulin. Another useful heterologous sequence is a functional domain which can substitute for a functional domain of mGluR5. For example, a chimeric receptor could be constructed in forder to functionally isolate aspects of mGluR5. See, for example, WO 97/05252.

In another embodiment, mGluR5 polypeptides can be conjugated with carrier or label. For example, in instances where the mGluR5 polypeptide is correctly configured so as to provide a binding site, but is too small to be immunogenic, the polypeptide may be linked to a suitable carrier. A number of techniques for obtaining such linkage are known in the art and need not be described in detail herein. Any carrier can be used which does not itself induce the production of antibodies harmful to the host. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins; polysaccharides, such as latex functionalized sepharose, agarose, cellulose, cellulose beads and the like; polymeric amino acids, such as polyglutamic acid, polylysine, and the like; amino acid copolymers; and inactive virus particles or attenuated bacteria, such as Salmonella. Especially useful protein substrates are serum albumins, keyhole limpet hemacyanin, immunoglobulin molecules, thyroglobulin, ovalbumin, tetanus toxoid, and other proteins well known to those of skill in the art. Labels are known in the art and are described herein.

Compositions containing mGluR5 polypeptides are also encompassed by this invention. When these compositions are to be used pharmaceutically, they are combined

with a pharmaceutically acceptable excipient. The compositions may alternatively or additionally contain a physiologically acceptable medium, examples of which have been provided above.

mGluR5 polypeptides of the invention can be identified and/or characterized in a number of ways. For example, an mGluR5 polypeptide can be tested for its ability to bind to, for instance, another protein (such as an antibody or ligand, such as glutamate). Alternatively, mGluR5 polypeptide(s) can be tested for its ability to elicit an immune response, whether humoral or cellular. It is understood that only one of these properties need be present in order for a polypeptide to come within this invention, although more than one of these properties may be present.

The ability of an mGluR5 polypeptide to bind (i.e., interact with) another protein can be assessed using standard techniques in the art. Binding of an mGluR5 polypeptide to an antibody may be assessed, for example, by RIA (i.e., by reacting radiolabeled mGluR5 polypeptide with an antibody that is coated on microtiter plates). In another procedure, binding to an antibody is determined by competitive immunoassay. For example, a fragment is tested for its ability to interfere with the binding between the antibody and another polypeptide known to bind to the antibody. This assay may be conducted by labeling one of the components (i.e., antibody or polypeptide known to bind to the antibody), and optionally immobilizing the other member of the binding pair on a solid support for ease of separation. The test fragment is incubated with labeled region, and then the mixture is presented to the immobilized target to determine if the test fragment is able to inhibit binding.

In the case of testing whether the mGluR5 polypeptide binds to another protein, for instance, a ligand, assays to detect binding are known in the art and need not be described in detail herein. For instance, a protein is immobilized on a suitable column. Extracts or solutions containing the test mGluR5 polypeptide are then run through the column, and a determination is made whether the mGluR5 polypeptide was retained on the column. Conversely, the mGluR5 polypeptides can be immobilized on a column and cell extracts or lysates are allowed to run through the column. Alternatively, the two hybrid technique can be used to identify polypeptides that interact with mGluR5 polypeptides (as well as, the cDNAs that encode such polypeptides) and to test such interactions. Brent et al. U.S. Pat. No. 5,580,736.

For characterizing an mGluR5 polypeptide for its ability to elicit an immune response (whether humoral or cellular) in an individual, standard assays exist in the art. For instance, the ability of an mGluR5 polypeptide to generate a humoral response can be determined by testing for the presence of an antibody that binds to the mGluR5 polypeptide(s) after administration of the mGluR5 polypeptide(s). It is understood that this antibody was not present, or was present in lower amounts, before administration of the mGluR5 polypeptide(s). Immunogenicity is preferably tested in individuals without a previous anti-mGluR5 response. Examples of suitable individual include, but are not limited to, mice, rats, rabbits, goats, monkeys and humans. For this test, an individual is administered an mGluR5 polypeptide(s). The amount per administration and the number of administrations will vary, depending on the individual. Presence of an antibody elicited in response to administration of an mGluR5 polypeptide(s) is determined by standard assays in the art, such as ELISA or RIA. mGluR5 polypeptide(s) may be further characterized by their ability to elicit an antibody that is cytotoxic, or to elicit an antibody that participates in an ADCC response using standard assays in the art.

A mGluR5 polypeptide can also be characterized by its ability to elicit a cellular immune response, using, for example, assays that detect proliferation of peripheral blood mononuclear cells (PBMs) incubated with an mGluR5 polynucleotide. Another way of detecting a cellular immune response is to test for T cell cytotoxicity (CTL) activity. Both of these responses are detected using standard assays in the art.

Other measurable activities may be used to characterize an mGluR5 polypeptide, such as membrane association, particular conformational change upon binding agonist(s) or antagonist(s), and ability to prevent or reduce signal transduction when expressed in a suitable host cell.

Preparation of polypeptides of this invention

The polypeptides of this invention can be made by procedures known in the art. The polypeptides can be produced by recombinant methods (i.e., single or fusion polypeptides) or by chemical synthesis. Polypeptides, especially shorter polypeptides up to about 50 amino acids, are conveniently made by chemical synthesis. Methods of chemical synthesis are known in the art and are commercially available. For example, a polypeptide could be produced by an automated polypeptide synthesizer employing the solid phase

method. Polypeptides can also be made by chemical synthesis using techniques known in the art.

Polypeptides can also be made by expression systems, using recombinant methods. The availability of polynucleotides encoding polypeptides permits the construction of expression vectors encoding intact (i.e., native) polypeptide, functionally equivalent fragments thereof, or recombinant forms. A polynucleotide encoding the desired polypeptide, whether in fused or mature form, and whether or not containing a signal sequence to permit secretion, may be ligated into expression vectors suitable for any convenient host. Both eukaryotic and prokaryotic host systems can be used. The polypeptide is then isolated from lysed cells or from the culture medium and purified to the extent needed for its intended use. Purification or isolation of the polypeptides expressed in host systems can be accomplished by any method known in the art. For example, cDNA encoding a polypeptide intact or a fragment thereof can be operatively linked to a suitable promoter, inserted into an expression vector, and transfected into a suitable host cell. The host cell is then cultured under conditions that allow transcription and translation to occur, and the desired polypeptide is recovered. Other controlling transcription or translation segments, such as signal sequences that direct the polypeptide to a specific cell compartment (i.e., for secretion), can also be used. Examples of prokaryotic host cells are known in the art and include, for example, *E. coli* and *B. subtilis*. Examples of eukaryotic host cells are known in the art and include yeast, avian, insect, plant, nematode, and animal cells such as COS7, HeLa, CHO and other mammalian cells.

When using an expression system to produce mGluR5 polypeptides, it is often preferable to construct a fusion protein that facilitates purification. Examples of components for these fusion proteins include, but are not limited to myc, HA, FLAG, His-6, glutathione S-transferase, maltose binding protein or the Fc portion of immunoglobulin. These methods are known in the art.

Alternatively, in vitro expression systems may also be used to produce mGluR5 polypeptides. A plasmid containing an mGluR5 polynucleotide, under the control of an appropriate promoter, can be transcribed and the resultant RNA translated in vitro through the use of commercially available reagents. Such methods can be used to produce relatively pure, although small amounts of the polypeptide and are known in the art.

Preferably, especially if used for diagnostic purposes, the polypeptides are at least partially purified from other cellular constituents. Preferably, the polypeptides are at least 50% pure. In this context, purity is calculated as a weight percent of the total protein content of the preparation. More preferably, the proteins are 50-75% pure. More highly purified polypeptides may also be obtained and are encompassed by the present invention. For clinical use, the polypeptides are preferably highly purified, at least about 80% pure, and free of pyrogens and other contaminants. Methods of protein purification are known in the art and are not described in detail herein.

Antibodies of the invention

Also provided by this invention are antibodies capable of specifically binding to mGluR5 polypeptide(s) of this invention. The antibodies can be useful for, for example, for detecting and characterizing mGluR5 polypeptides, as described above. Antibodies of this invention can also be used for purification and/or isolation of polypeptides described herein.

In one embodiment, the invention provides a purified antibody capable of specifically binding to a polypeptide of this invention. As noted in the definition of "antibody" above, this includes fragments of antibodies, such as Fab fragments. In another embodiment, a monoclonal antibody is provided that is capable of specifically binding to a polypeptide of this invention.

Laboratory methods for producing polyclonal antibodies and monoclonal antibodies, as well as deducing their corresponding nucleic acid sequences, are known in the art. For example, see Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York (1988) and Sambrook et al. (1989).

The antibodies of this invention may be polyclonal or monoclonal. Monoclonal antibodies of this invention can be biologically produced by introducing a polypeptide (or fragment of a polypeptide) of this invention into an animal, e.g., mouse or rat . The antibody producing cells in the animal are isolated and fused with myeloma cells or heteromyeloma cells to produce hybrid cells or hybridomas. Accordingly, the invention also includes hybridoma cells producing the monoclonal antibodies of this invention.

Particular isotypes of a monoclonal antibody can be prepared either directly by selecting from the initial fusion, or prepared secondarily, from a parental hybridoma

secreting a monoclonal antibody of different isotype by using the sib selection technique to isolate class switch variants using the procedure described in Steplewski et al. (1985) *Proc. Natl. Acad. Sci.* 82:8653 or Spira et al. (1984) *J. Immunol. Methods* 74:307.

Thus, using the polypeptide(s) of this invention or fragment(s) thereof, and well known methods, one of skill in the art can produce and screen the hybridoma cells and antibodies of this invention for antibodies having the ability to bind polypeptide(s) of this invention.

If a monoclonal antibody being tested binds with an mGluR5 polypeptide(s) of this invention, then the antibody being tested and the antibodies provided by the hybridomas of this invention are equivalent. It is also possible to determine without undue experimentation whether an antibody has the same specificity as a monoclonal antibody of this invention by determining whether the antibody being tested prevents a monoclonal antibody of this invention from binding the polypeptide(s) with which the monoclonal antibody is normally reactive. If the antibody being tested competes with the monoclonal antibody of the invention as shown by a decrease in binding by the monoclonal antibody of this invention, then it is likely that the two antibodies bind to the same or a closely related epitope. Alternatively, one can pre-incubate the monoclonal antibody of this invention with the polypeptide(s) with which it is normally reactive, and determine if the monoclonal antibody being tested is inhibited in its ability to bind the antigen. If the monoclonal antibody being tested is inhibited, then, in all likelihood, it has the same, or a closely related, epitopic specificity as the monoclonal antibody of this invention.

As noted above, this invention also provides biological active fragments of the polyclonal and monoclonal antibodies described above. These antibody fragments retain some ability to selectively bind with its antigen or immunogen. Examples of antibody fragments are known in the art and include, but are not limited to, CDR regions, Fab, Fab', $F(ab')_2$, F_V, and single chain methods. Methods of making these fragments are known in the art, see for example, Harlow and Lane, (1988).

The antibodies of this invention also can be modified to create chimeric antibodies and humanized antibodies (Oi et al. (1986) *BioTechniques* 4(3):214). Chimeric antibodies are those in which the various domains of the antibodies' heavy and light chains are coded for by DNA from more than one species.

The isolation of other hybridomas secreting monoclonal antibodies with the specificity of the monoclonal antibodies of the invention can also be accomplished by one skilled in the art by producing anti-idiotypic antibodies (Herlyn, et al. (1986) *Science*, 232:100). An anti-idiotypic antibody is an antibody which recognizes unique determinants present on the monoclonal antibody produced by the hybridoma of interest. These determinants are located in the hypervariable region of the antibody. It is this region which binds to a given epitope and, thus, it is responsible for the specificity of the antibody. The anti-idiotypic antibody can be prepared by immunizing an animal with the monoclonal antibody of interest. The animal immunized will recognize and respond to the idiotypic determinants of the immunizing antibody by producing an antibody to these idiotypic determinants. By using the anti-idiotypic antibodies of the second animal, which are specific for the monoclonal antibodies produced by a single hybridoma which was used to immunize the second animal, it is now possible to identify other clones with similar idiotypes as the antibody of the hybridoma used for immunization.

Idiotypic identity between monoclonal antibodies of two hybridomas demonstrates that the two monoclonal antibodies are the same with respect to their recognition of the same epitopic determinant. Thus, by using antibodies to the epitopic determinants on a monoclonal antibody it is possible to identify other hybridomas expressing monoclonal antibodies of the same epitopic specificity.

It is also possible to use the anti-idiotype technology to produce monoclonal antibodies which mimic an epitope. For example, an anti-idiotypic monoclonal antibody made to a first monoclonal antibody will have a binding domain in the hypervariable region which is the mirror image of the epitope bound by the first monoclonal antibody. Thus, in this instance, the anti-idiotypic monoclonal antibody could be used for immunization for production of these antibodies.

The antibodies of this invention can be linked (i.e., conjugated) to a detectable agent or a hapten. The complex is useful to detect the polypeptide(s) (or polypeptide fragments) to which the antibody specifically binds in a sample, using standard immunochemical techniques such as immunohistochemistry as described by Harlow and Lane (1988). *supra*. Examples of types of immunoassays which can utilize monoclonal antibodies of the invention are competitive and non-competitive immunoassays in either a direct or indirect format. Examples of such immunoassays are the enzyme linked immunoassay (ELISA)

radioimmunoassay (RIA) and the sandwich (immunometric) assay. Detection of using the monoclonal antibodies of the invention can be done by utilizing immunoassays which are run in either the forward, reverse, or simultaneous modes, including immunohistochemical assays on physiological samples. Those of skill in the art will know, or can readily discern, other immunoassay formats without undue experimentation.

Another technique which may also result in greater sensitivity consists of coupling the antibodies to low molecular weight haptens. These haptens can then be specifically detected by means of a second reaction. For example, it is common to use such haptens as biotin, which reacts avidin, or dinitropherryl, pyridoxal, and fluorescein, which can react with specific anti-hapten antibodies. See Harlow and Lane (1988) *supra*.

The monoclonal antibodies of the invention can be bound to many different carriers. Thus, this invention also provides compositions containing antibodies and a carrier. Carriers can be active and/or inert. Examples of well-known carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, agaroses and magnetite. The nature of the carrier can be either soluble or insoluble for purposes of the invention. Those skilled in the art will know of other suitable carriers for binding monoclonal antibodies, or will be able to ascertain such, using routine experimentation.

There are many different labels and methods of labeling known to those of ordinary skill in the art. Examples of the types of labels which can be used in the present invention include enzymes, radioisotopes, fluorescent compounds, colloidal metals, chemiluminescent compounds, and bioluminescent compounds. Those of ordinary skill in the art will know of other suitable labels for binding to the monoclonal antibody, or will be able to ascertain such, using routine experimentation. Furthermore, the binding of these labels to the monoclonal antibody of the invention can be done using standard techniques common to those of ordinary skill in the art.

For purposes of the invention, mGluR5 polypeptides of this invention may be detected by the monoclonal antibodies of the invention by their ability to bind these antibodies.

Compositions containing the antibodies, fragments thereof or cell lines which produce the antibodies, are encompassed by this invention. When these compositions are to be used pharmaceutically, they are combined with a pharmaceutically acceptable

excipient. The compositions may alternatively or additionally contain a physiologically acceptable medium, examples of which have been provided above.

The following examples are provided to illustrate but not limit the invention.

EXAMPLES

Example 1: Differential effect of CNS depressant propofol on LS and SS mice

Methods

Experimental protocols were approved by our animal care committee. Male LS and SS mice were provided by the Institute for Behavioral Genetics, University of Colorado, Boulder. They were of the 54th generation (25th generation of selection), were weaned at 25d, and were drug tested at 60-90 d.

Right-sided jugular venous cannulas were implanted under intraperitoneal pentobarbital (60 mg/kg)-chloral hydrate (125 mg/kg) anesthesia on Day 1. Mice were allowed to rest for 2 d and propofol testing took place on Day 3. Propofol formulated in Intralipid (10 mg/kg) was obtained from Stuart Chemicals (Wilmington, DE) and administered intravenously via the jugular cannula. Injected volumes ranged from 30 to 70 µl and were given over 30-45 s by Hamilton syringe. Doses appropriate for loss of righting reflex ranged from 10 to 35 mg/kg. Saline flush was administered after propofol to bring the total injectate to 100 µl. The loss of righting reflex was immediate in all cases. Animals were placed on their backs in V-shaped Plexiglas troughs and judged awake upon regaining the righting reflex (ability to turn over three times in 1 min).

Propofol levels at awakening were determined in brain cortex and plasma samples. The method of intravenous propofol dosing was simplified by administering 20 mg/kg propofol via the retroorbital venous sinus using a 28-gauge needle and Hamilton syringe. Injected volumes ranged from 40 to 50 µL and were given over 30 s. At awakening, cortical brain and body cavity samples were collected. Blood samples were pooled and centrifuged at 3000*g* for 15 min; plasma was collected and stored at 4° C until analysis by gas chromatography. Brain samples were weighed and stored until analysis by gas chromatography. The method of Yu and Liau was used for determination of brain and plasma levels by gas chromatographic analysis. *J. Chromatog.* (1993) 615:77-81. 2-*sec*-Butylphenol was used as an internal standard.

$^{36}Cl^-$ uptake was performed using freshly prepared mouse brain microsacs as described previously. Allan and Harris (1986) *Life Sci.* 39:2005-2015. Brain cortices were manually homogenized in a glass Teflon homogenizer. Homogenizing solution was 4.5 mL of ice-cold assay buffer containing (mM) NaCl 145, KCl 5, $MgCl_2$ 1, d-glucose 10, $CaCl_2$ 1, and HEPES 10, adjusted to pH 7.5 with Tris base. The homogenate was centrifuged at 900*g* for 15 min, the supernatant decanted, and the pellet resuspended in 8 mL of assay buffer. It was then centrifuged again at 900*g* for 15 min. The final pellet was suspended in 7 mL of assay buffer. Protein concentrations were determined by the method of Lowery et al. *J. Biol. Chem.* (1951) 193:265-275.

Propofol stimulated $^{36}Cl^-$ uptake was maximized in the following way: aliquots of microsacs (200 μL) were incubated in a shaking water bath at 34°C for 5 min. $^{36}Cl^-$ uptake was initiated by addition of 200 uL of a solution containing $^{36}Cl^-$ (2 μCi/mL). Muscimol (2 μM) and propofol were added in the $^{36}Cl^-$ solution. $^{36}Cl^-$ influx was terminated after a 3-s incubation period by addition of 4 mL of ice-cold buffer and rapid filtration under vacuum onto Whatman GF/C glass microfiber filters. Filters were washed with an additional 8 mL of cold buffer. Radioactivity retained on filters was determined by liquid scintillation spectrometry. Control values were obtained using only Intralipid in the preparation. Blank values (no tissue in preparation) were subtracted from all final values. Muscimol-dependent uptake was defined as the amount of $^{36}Cl^-$ taken up with agonist present in the medium minus the amount of $^{36}Cl^-$ taken up while agonist was absent (muscimol-independent uptake). Propofol-stimulated uptake was calculated as the uptake in the presence of propofol plus muscimol minus the uptake in the presence of muscimol alone.

Student's unpaired *t*-test was used for comparisons of mean sleep times and drug levels. The $^{36}Cl^-$ uptake of brain microsacs was compared at several different propofol concentrations using analysis of variance. Statistical significance was set at $P < 0.05$.

Results

Propofol produced dose-dependent increases in sleep times for both LS and SS mice as shown in Figure 1. The LS mice clearly displayed increased sensitivity to propofol, as manifested by longer sleep times, compared with SS mice at every dose. Plasma and cortical brain levels of propofol at awakening were measured and found to be different in the LS and SS mice ($P < 0.0001$). The LS mice had threefold greater plasma levels and twofold greater brain levels than did the SS mice as shown in Table 1. Propofol plasma

levels are expressed in μg/mL plasma ± SD. Propofol brain levels are expressed in μg/g tissue ± SD. A propofol dose of 20 mg/kg was given by retroorbital sinus injection. For both comparative measurements, there was a $P < 0.0001$ significant difference between means of LS and SS lines.

Table 1. Propofol Plasma and Brain Levels at Awakening in Long Sleep (LS) and Short Sleep (SS) Mice

Tissue	LS (*n*)	SS (*n*)
Plasma	4.52 ± 0.91 (9)	13.3 ± 1.15 (9)
Brain	8.05 ± 0.9 (13)	16.0 ± 2.4 (15)

The method of propofol dosing was changed from the jugular route (JUG) via the previously implanted catheter to retroorbital venous sinus (ROS) injections using an intravenous dose of 20 mg/kg. We found that the JUG method added several days to a single experiment, produced significant mortality associated with the surgical procedure, and added a prior exposure to the general anesthetics needed to implant the jugular catheter. The ROS method circumvented all these problems. Figure 2 shows the results comparing each of these administrations in SS and LS mice. Although there was a statistically significant difference in sleep times ($P < 0.02$) in the SS line using JUG dosing versus ROS dosing, the difference was only 1.2 min. (Figure 2). A similar but nonsignificant difference of 1.4 min. was observed in the LS line. The difference in sleep time between the SS and LS lines remained very significant ($P < 0.0005$). For both lines, the sleep times were somewhat shorter using JUG dosing as compared to ROS dosing. SS sleep time means ±SD were 2.24 ±0.51 min. (JUG) and 3.45 ± 1.02 min. (ROS). LS sleep time means were 4.41 ± 1.29 min. (JUG) and 5.66 ± 1.48 min. (ROS).

We then tested effects of increased propofol on muscimol-stimulated $^{36}Cl^-$ uptake in LS and SS mice. The results are shown in Figure 3. Propofol produced a concentration-dependent enhancement of muscimol-stimulated $^{36}Cl^-$ uptake by membrane vesicles (microsacs) prepared from mouse cortical brain. (Figure 3) However, there was no significant difference noted between LS and SS brain preparations in the increase in muscimol-stimulated $^{36}Cl^-$ uptake at any of the propofol concentrations. Significant

differences in mean $^{36}Cl^-$ uptake were observed between doses by differed by 100-fold concentrations. Propofol concentrations ranged from 0.50 to 50 μmol/L (corresponding to brain levels of 0.32-31.6 μg/mL). Propofol did not affect basal uptake of $^{36}Cl^-$ except at the highest concentration tested (50 μmol/L), which increased basal uptake by about 25%.

Example 2: Mapping CNS depressant sensitivity

In the absence of a biochemical hypothesis to explain the LS and SS difference, we used a positional cloning strategy to determine the location of the locus (or loci) contributing to this activity.

Materials and Methods

Origin of Mice

ILS, ISS, and LSXSS RI mice were provided by the Institute for Behavioral Genetics, University of Colorado at Boulder. ILS and ISS have been derived by 20 rounds of brother-sister mating from the LS and SS selected lines; the LSXSS RI set was similarly derived after reciprocal intercrosses between LS and SS, prior to the LS and SS having been inbred. Market et al. (1996); DeFries et al. (1989) *Alcohol Clin Exp Res* 13:196-200. Thus the progenitors of the LSXSS RIs and the F_2 intercross are not the same stock and could have some differences affecting sensitivity to propofol. Animals were bred in a specific-pathogen free facility, weaned at about 25 days and housed 1-4 per cage with like-sex littermates on a 12-hour light/dark cycle. Experimental protocols were approved by animal care and use committees at the University of Colorado at Boulder and University of Colorado Health Sciences Center at Denver. For the LSXSS RIs, between 5 and 15 mice of each sex were tested per strain. For the ILS x ISS F_2s, 194 mice were tested for LORR and 164 subsequently genotyped. C57BL/6J coisogenic mice having the spontaneous albino mutation c^{2J} were obtained from Jackson Laboratory, Bar Harbor, Maine.

Phenotypic Assessments

Propofol (10 mg/ml in Intralipid, Stuart Chemicals) injections were performed when mice were 60 to 94 days of age. Dosages of 20 mg/kg (volumes ranging from 30-70 μl) were administered by injection into the retroorbital venous sinus using a 26-gauge, 3/8-inch needle and Hamilton syringe. Sensitivity to propofol was assessed using the duration of loss of righting reflex (LORR). Simpson et al. (1996). Upon injection, mice

immediately lost consciousness and were placed on their back in V-shaped Plexiglass troughs until recovery of righting reflex. Animals were judged to have regained the righting reflex when they turned over three times in less than a minute. Temperature was controlled at 24-26°C. Full recovery was quick after regaining consciousness, as typically observed in recovery from propofol anesthesia in humans; after the first righting, there was almost no delay in the mouse's ability to right a second and third time. The RI and F_2 mice were naive to any previous injections and were injected only once with propofol. LORRs of the (ISS *c/c* x C57BL/6 c^{2J}/C) F_1 mice represent the mean of two propofol injections spaced seven days apart whose means were not significantly different ($p = 0.36$, ANOVA two-tailed significance). The brain levels of propofol at awakening (BLA) for the RIs were determined in pooled samples of brain cortex using gas chromatography, as previously described by Simpson and Blednov (1996).

Marker Assessments

DNA was extracted and purified from spleen, kidney, or liver using a Super-Quik Gene DNA Isolation Kit (AGTC Research). The *Mit* SSLP primers were obtained from Research Genetics and the PCR was carried out as previously described. Markel et al. (1994) *Mam. Genome* 5:199-202. Coat color was scored by eye on living mice; pigmented RI mice were homozygous C/C and albino RI mice were homozygous *c/c*. Jackson et al. (1984) *Proc. Natl. Acad. Sci.* 87:7010-14. (This locus was renamed Tyr because it encodes tyrosinase, an enzyme responsible for pigmentation in mammals. See Beerman et al. (1990) *EMBO J.* 9:2819-2826; Yokoyama et al. (1990) *Nucl. Acids Res.* 18:7293-7298).)

Statistical Analyses

Broad-sense heritability involves genetic effects arising from dominant, additive and between-gene interactions while narrow-sense heritability involves only additive effects; thus broad-sense heritability is more inclusive and should be a larger fraction than narrow sense. Narrow-sense heritability in the RIs was estimated as $100\% \times 0.5\ \sigma^2_B/(\sigma^2_W + 0.5\sigma^2_B)$, in which σ^2_B is the between-strain component of variance and σ^2_W is the within-strain component of variance. Falconer (1981) *Quantitative Genetics*, Ronald, New York. Broad sense heritability was estimated from the ILS by ISS F_2s as $100\% \times (V_{F2} - V_{F1})/V_{F2}$, where V_{F2} is the variance of the F_2 population and V_{F1} represents an average variance calculated from the F_1, ILS, and ISS variances. The QTL maps in Figures 2 and 3 were anchored using map positions from the Mouse Genome Database (MGD) map

(*http://www informatics.jax.org/ mgd.html*) of 44.0 for *c* and D7*Mit*31, whereas the positions of other SSLP markers (Copeland et al. (1993) *Science* 262:57-66; Dietrich et al. (1994) *Nature Genet.* 7:220-225; Dietrich et al. (1996) *Nature* 380:149-152) relative to these were based on the Whitehead Institute/MIT Center for Genome Research (WI/MIT CGR) map (http://www.genome.wi.mit.edu), which is the more recently updated and more accurate map regarding the relative position of SSLP markers in this region. Rikke et al. (1997) *Genetics* 147:787-799. The RI map was constructed using an interval-mapping procedure previously described by Markel et al. (1996) *Behav. Genet.* 26:447-458. The F_2 linkage map was constructed using Mapmaker 3.0 and QTL mapping was done using Mapmaker/QTL Version 1.1.

Results

Heritability of Propofol Sensitivity

The results are shown in Table 2. For each population or cross, the female parent is listed first. There was a significant difference in sensitivity to propofol between the LS and SS lines ($p << .001$; Table 2, lines 1, 2), as described in Example 1 (and previously reported by Simpson et al. (1996)); their inbred derivatives, ILS (LORR of 7.4 minutes) and ISS (LORR of 3.8 minutes), also are differentially sensitive to propofol ($p << .001$; Table 2, lines 3, 4). The ILS and ISS strains were also significantly different from the two F_1 reciprocal populations with respect to propofol sensitivity. The F_1 showed no statistically significant differences (Table 2, lines 5, 6; $p = 0.08$, one-tailed) resulting from the maternal parent, thus indicating a lack of maternal effects for sensitivity to propofol. The mean LORR of the F_1 was almost intermediate between the two parents but showed slight dominance of the LS allele.

Table 2. Duration of loss of righting reflex (LORR) after propofol injection in various mouse populations.

Population	Trait	Total N	Mean ± SEM (minutes)	Variance (minutes)
1. LS	LORR	19	5.7 ± 0.3	1.8
2. SS	LORR	16	3.5 ± 0.2	0.88
3. ILS	LORR	19	7.4 ± 0.4	2.7
4. ISS	LORR	37	3.8 ± 0.2	2.2
5. F1, ILS x ISS	LORR	17	6.5 ± 0.5	3.8
6. F1, ISS x ILS	LORR	8	5.3 ± 0.7	3.4
7 F_1 combined	LORR	25	6.1 ± 0.4	3.7
F_2 ILS/ISS x ILS/ISS	LORR			
8. pigmented		32	6.7 ± 0.5	9.6
9. albino		18	6.1 ± 0.6	6.7
F_2 ILS/ISS x ILS/ISS	LORR			
10. pigmented		35	6.8 ± 0.4	7.0
11. albino		8	4.3 ± 0.8	5.3
F_2 ISS/ILS x ILS/ISS	LORR			
12. pigmented		36	6.7 ± 0.3	2.9
13. albino		18	6.2 ± 0.7	8.0
F_2 ISS/ILS x ISS/ILS	LORR			
14. pigmented		32	7.1 ± 0.6	9.8
15. albino		15	5.7± 0.6	4.6
F_2 All	LORR	194	6.5 ± 0.2	6.9
16. pigmented		135	6.8 ± 0.2	7.1
17. albino		59	5.8 ± 0.3	6.5
18. N_2 ILS x ILS/ISS	LORR	27	6.6 ± 0.4	4.9
19. N_2 ISS/ILS x ILS	LORR	28	6.0 ± 0.5	6.1
N_2 ISS x ILS/ISS				
20. pigmented	LORR	19	6.2 ± 0.6	7.8
21. albino	LORR	6	5.2 ± 0.6	2.2
22. total	LORR	25	6.0 ± 0.5	6.5
N_2 ILS/ISS x ISS				
23. pigmented	LORR	15	7.6 ± 0.6	4.7
24. albino	LORR	14	6.7 ± 0.5	3.2
25. total	LORR	29	7.1 ± 0.4	4.1
			(μg/g brain)	
LS	BLA	11	8.2 ± 0.2	0.81
SS	BLA	10	16.1 ± 0.6	5.8

Estimates of heritability show what fraction of the variation among individuals is due to genetic effects. Falconer (1981). We have estimated these heritabilities for differential sensitivity to propofol using two distinct approaches. First, an estimate of broad-sense heritability (h^2) can be obtained by assessing the variation in length of LORR after propofol injection in the inbred parental strains (ILS and ISS) and in the F_2 reciprocal

cross (Table 2). An F2 population derived from an ILS by ISS and the reciprocal intercross had a mean not different from the F_1. Comparing the F_2 variance with the average variance of the F_1 and the two parental strains we derived an estimate of broad-sense heritability of 55%. Falconer (1981).

A second estimate can be obtained by dividing 1/2 the variance among the 24 RI strain means by 1/2 the inter-strain variance plus the variance within the 24 RIs. Falconer (1981). Using this method, which estimates narrow-sense heritability, we found that about 32% of the variance is genetic. Broad-sense heritability includes estimates of dominance and interactions among loci while narrow-sense heritability only estimates additive effects; these results thus suggest that there may be a substantial genetic component resulting from epistatic (between different genes) interactions, but both methods suggest a large overall genetic component. An estimate of BLA heritability could not be obtained because of an inability to collect BLA on individual mice due to the limited sensitivity of the assay.

Location of a QTL for Propofol Sensitivity

The LS and SS selected lines are almost completely differentiated at the albino locus (*Tyr*) with all SS mice being albino and most of the LS mice being pigmented. Surprisingly, when we examined the LSXSS RIs we found that the eleven albino RI strains averaged significantly shorter LORRs in response to propofol than did the 13 pigmented RI strains ($p < .0002$), with no difference between males and females as shown in Fig. 4A. The difference in LORR between males and females was not significant; whereas, the LORR difference between pigmented (black bars, 4.8 ± SEM) and albino RI strains (white bars, 9.3 SEM) was highly significant for both males and females ($p << 0.0001$, Student's *t* test two-tailed significance). "n" refers to the number of RI strains. The LORR for each strain was determined from the mean of 5-15 animals of each sex, and the BLA for each strain was determined from the mean of at least 3 animals. The difference in propofol BLA between pigmented (8.4 ± 1.2 μg/gm) and albino (17.2 ± 1.9 μg/gm) strains is highly significant ($p << 0.0001$, Student's *t* test two-tailed significance). The pigmented RI strains (black bars) used for propofol BLA determination were 2, 3, 6, 7, 8, 9, 17, 30 and 36; and the albino RI strains (white bars) used were 4, 10, 18, 20, 23, 25, and 33. Male and female data are combined for propofol BLA. The correlation between propofol BLA and LORR for the 16 RI strains for which both measures could be obtained was -0.85 ($p < 0.001$). RI 32 was excluded from this analysis because it was still segregating for the albino mutation.

The albino strains also had very significantly higher BLA ($p << .0001$; Fig 1A), suggesting the difference in LORR is due to differential CNS sensitivity and not pharmacokinetic differences. The correlation coefficient between BLA and LORR was 0.85 and highly significant ($p < 0.001$). The LORRs of the albino strains were non-overlapping with the LORRs of the pigmented strains, demonstrating a major effect of a QTL linked to or at the *Tyr* locus.

This QTL was mapped using the regression approach described by Haley and Knott, which we previously used in mapping QTLs for alcohol sensitivity. We assessed the genotypes of seven additional SSLP markers spanning 50 centiMorgan (cM) of chromosome 7, obtaining LOD scores of 19 and 21 for males and females, respectively. The combined LOD score for males and females was about 28 (Fig. 5), far exceeding the value of 3.3 recommended by Lander and Kruglyak for significant linkage in a whole-genome scan. This assignment was highly specific and positioned the QTL to within 2.5 cM of *Tyr* with 2-LOD support, indicating the probability of the QTL being within this region is greater than 0.99. The peak LOD score was at the *Tyr* gene itself. This QTL, which we have called *Lorp1* (*Loss Of Righting* due to Propofol), explained about 80% of the genetic variance between LS-like and SS-like RI strains, indicating this is the major gene specifying differential propofol sensitivity between LS and SS. BLA also mapped to this region, peaking at the *Tyr* locus, although with significantly reduced LOD score, not passing the suggested level of 3.4 needed to assure that the QTL is real. With respect to Fig. 7, ISS mice have the classical albino mutation, whereas the c^{2j} albino mutation arose spontaneously in C57BL/6 at Jackson Laboratory. Therefore, F_1 mice are identical at all loci except the albino locus, pigmented mice (black bars) are *c/C* and albino mice (white bars) are c/c^{2J}. LORRs of females are significantly shorter than those of males by 1.3 minutes ($p = 0.002$, ANOVA two-tailed significance). However, LORRs of albinos are not significantly shorter than those of pigmented mice ($p = 0.35$, ANOVA one-tailed significance).

Confirmation of Linkage

Although confirmation of linkage after such an astronomical LOD score is not needed, the map location of *Lorp1* was confirmed independently using a panel of 164 F_2 mice derived from intercrosses of ILS and ISS. These F_2s were genotyped with 16 SSLP markers from murine chromosome 7 that included all of the informative SSLP markers

flanking *Tyr* for about 10 cM (WI/MIT CGR map and Rikke et al., submitted). A peak LOD score of 3.9 was obtained between markers *D7Mit31*, which is within *Tyr*, and *D7Mit123*, which is less than 2 cM distal to *Tyr* (Fig 6). We found that F_2 homozygotes for the SS allele of *D7Mit123*, slept 5.7 minutes (SEM = 0.3), heterozygotes slept 6.5 minutes (SEM = 0.3), and LS homozygotes slept 8.0 minutes (SEM = 0.4). The 2.3 minute difference in propofol-induced LORR between mice homozygous for the S allele at *D7Mit123* and those homozygous for the L allele was consistent with the 2.2 minute difference previously observed between LS and SS, suggesting that all of the genetic difference between LS and SS for propofol sensitivity has been captured in the ILS and ISS strains and that most or all of the genetic difference resides in *Lorp1*.

Location of a QTL for Ethanol, Enflurane, Isoflurane and Etomidate

The QTL location experiments described above were repeated for CNS depressants ethanol (4.1 g/kg), isoflurane (2.0 g/kg), enflurane (5.0 g/kg) and etomidate (20 mg/kg). The behavioral measure of sensitivity was LORR (as was used in the propofol studies, above). The results (including those for propofol) are shown in Figs. 8-10. The peak LOD scores are at or very close to *Tyr* for each CNS depressant, just was with propofol. However, the 2-LOD support interval for each is much broader than 2.5 cM as reported for propofol (on the order of 25-30 cM). The results indicate that the genetic locus associated with differential CNS sensitivity is linked to the *Tyr* locus.

Example 3: A 250 kb fragment in yeast artificial chromosome YRT2 confers increased sensitivity to propofol

Transgenic mice harboring YAC YRT2 (covering the mouse tyrosinase locus as isolated from a YAC library of C3H mouse DNA) were kindly provided by Schütz. Schedl et al. (1993) *Nature* 362:258-261. YRT2, the resulting 250 kb construct, contains 80 kb of the tyrosinase coding region (Ruppert, S. et al. *EMBO J.* 7:2715-2722, (1988)), 155 kb of upstream sequences and 15 kb of vector DNA (Fig. 11).

The transgenic mice were originally created in outbred albino mouse strain NMRI. These mice were crossed with albino ISS strain. Progeny from crossing two different strains were denoted F_1S. Thus, F_1S mice were 50% NMRI and 50% ISS. The F_1S mice

were then backcrossed with ISS, and these progeny were denoted N_2S (which were 25% NMRI and 75% ISS). For both F_1S and N_2S, mice having the transgene were pigmented, whereas non-transgenic littermates were albinos.

We tested three different YRT2 lines for these experiments. Transgenes insert at random sites in the genome; therefore, it was presumed that each line had its transgenic insert at a different site. Transgenes also tend to integrate as tandemly repeated copies of the original transgenic insert. Schedl et al. determined that one of the transgenic lines had one copy of the transgenic region (denoted Tg1), another line had two copies (Tg2), and another line had eight copies (Tg8). More copies of the transgene generally implies greater expression levels of the transgene, which could explain the greater different seen with the Tg8 mice. However, expression levels may also depend on where in the genome the DNA has inserted.

Sleep times of transgenic mice were compared to sleep times of non-transgenic littermates. All mice (60 to 90 days old) received 20 mg/kg propofol The results are shown in Fig. 12. N indicates the number of mice tested. For each comparison, the difference was statistically significant. The transgenic mice showed longer sleep time (i.e., increased sensitivity) than non-transgenic mice. In the first group, transgenic mice slept an average of 0.7 minutes longer than non-transgenic mice ($p= 0.01$). In the second group, transgenic mice slept an average of 0.6 minutes longer than non-transgenic mice ($p=0.056$). In the third group, transgenic mice slept an average of 1.5 minutes longer than non-transgenic mice ($p=0.03$). Sleep-time sensitivity was independent of transgene insertion site, indicating that the locus *Lorp1* is within the transgenic region, i.e., within the 250 kb YAC insert.

Example 4: Analysis of sequences on YRT2, including mGluR5 sequences

We analyzed sequences contained in YRT2 in an effort to determine which sequences of the 250 kb insert may be particularly associated with differential CNS depressant activity. To accomplish this, we first isolated a BAC clone as follows. A mouse BAC genomic library was obtained from a commercial source (Research Genetics). The library was screened by PCR using primer pairs corresponding to known genetic markers that were known to reside on the 250 kb insert of YRT2. A positive BAC clone

was identified.. To convert the BAC into a small insert sequence-ready library, we followed the scheme depicted in Fig. 13. BAC DNA was sheared using a sonicator and fragments of about 1.5 kb were size selected on agarose gels and isolated. The size-selected fragments were blunt end ligated into a pBLUESCRIPT™ vector which had been SmaI linearized and dephosphorylated. The ligated vectors containing the size-selected fragments were electroporated into XLI-BLUE™ electrocompetent cells. Blue colonies were picked, and the inserts sequenced using fluorescence (Applied Biosystems AB1373A).

Once a sequence had been obtained, it was compared to sequences in GenBank using a BLAST searches. The comparison revealed that some sequences were homologous to human and rat mGluR5, with rat mGluR5 exhibiting the highest homology (SEQ ID NOS:5 and 6). A more extensive sequence of the mGluR5 gene was obtained by sequencing this region using DNA amplified from total brain cDNA from LS and SS mice.

The polynucleotide sequence of a coding region of mouse mGluR5 is shown in SEQ ID NO:1. A polypeptide sequence for mouse mGluR5 is shown in SEQ ID NO:2. A comparison of over 90% of the amino acid sequence for mGluR5 in LS and SS mice showed no differences in either the amino acid or polynucleotide sequences.

Example 5: Glutamate receptor modulation affects propofol sensitivity

LS and SS mice receiving propofol and agonist ACPD

LS and SS mice were tested for the effect of ACPD, a group I metabotropic agonist, on propofol sleep time. Five LS (sensitive) mice (60 to 90 days old) and three SS (resistant) mice (60 to 90 days old) were used for this experiment. Intracerebroventricular catheters were implanted under pentobarbital (PB, 80 mg/kg) chloral hydrate (CH, 120 mg/kg) anesthesia, administered by intraperitoneal injection. After a 3 day recovery, propofol was administered at 20 mg/kg by retroorbital sinus injection. 5 µl of agonist ACPD (Tocris-Cookson) was administered into the lateral ventricles via the implanted catheter. After 30 minutes, the propofol sleep time was assessed. Three LS and three SS control mice

received 5 µl saline instead of agonist. Sleep time, measured as LORR, was recorded for test mice and controls. After sleep times were measured, the mice were sacrificed and the brains dissected to confirm the proper placement of the ventricular catheter. In all animals, the catheter was in proper position.

The results are tabulated in Table 4. SS (resistant) mice receiving agonist displayed a greater degree of increased resistance (when compared to SS controls) than LS (sensitive) mice receiving agonist. SS mice slept an average of 2.70 minutes shorter with agonist than without agonist, while LS mice slept an average of 0.67 minutes shorter with agonist than without agonist. These results are consistent with the hypothesis that the mGluR system (particularly those in Class I) is more sensitive to stimulation by agonist in SS (resistant) mice than LS mice, and thus may at least partially account for the difference in responsiveness between LS and SS mice.

Transgenic mice receiving agonist

Transgenic mice as described in Example 3 received approximately 40 µl propofol (20 mg/ml) and 5 µl agonist ACPD (62.5 pmol/5 µl). Sleep times were measured as described. The results are shown in Table 3. Transgenic mice slept an average of 3.97 minutes, while control (non-transgenic littermates) mice slept an average of 0.83 minutes. The average (mean) difference was 3.14 minutes, with a p value of 0.028523.

Example 6: Screening Using YRT2 Sequences

The following example describes cell-based and in vivo screening assays which use mGluR5 polynucleotide sequences found within the YRT2 sequence, but any analogous sequences (e.g., transcriptional control regions, polypeptide-encoding sequences) found within the YRT2 sequence may also be used in the following screens and assays.

In an example to identify such agents, an mGluR5-green fluorescent protein (GFP) hybrid gene is assembled in which the expression of GFP is under the control of the mGluR5 transcriptional control region and this hybrid gene is stably introduced into a mammalian cell line (e.g., CHO cells or a cell line with neuronal characteristics).

To screen for agents that modulate the expression of the mGluR5 gene (and may thus effect CNS depressant activity), cells that carry the mGluR5-GFP hybrid gene are grown in standard media and an agent to be tested is added (control cells receive no agent). After a suitable time, the cell suspensions are checked for GFP fluorescence. These assays

may be rapidly and conveniently performed in microtiter plates, using a small amount of media and agent to be tested in each well with the cells. Agents are identified by their ability to increase or decrease the amount of GFP fluorescence relative to that of the control cells. A relative increase in GFP fluoresence in cells grown in the presence of the agent may indicate an activation of expression driven from the mGluR5 regulatory elements by the agent. A relative decrease in GFP fluoresence in cells grown in the presence of the agent may indicate a suppression of expression from said regulatory elements. This mGluR5-GFP hybrid gene assay is performed in the presence of a CNS depressant to screen for agents which may modulate the depressant activity on mGluR5 expression, if any.

In another example to identify agents which exhibit CNS depressant activity or which modulate the activity of a CNS depressant, sequences encoding an mGluR5 polypeptide are expressed in cells (e.g., CHO cells or a neuronal cell line) under control of a promoter active in the particular cell type chosen. The promoter may be the endogenous mGluR5 promoter or a heterologous (including an inducible) promoter.

To screen for agents that modulate the activity of the mGluR5 polypeptide (and may thus effect CNS depressant activity), cells that express the mGluR5 polypeptide are grown in appropriate media and an agent to be tested is added (control cells receive no agent). After a suitable time, mGluR5 activity in the cells is measured through, for example, the determination of intracellular calcium release (using, for example, a fluorescent calcium indicator such as fura-2). Agents are identified by their ability to increase or decrease the activity of mGluR5 relative to that in the control cells. This mGluR5 activity assay is performed in the presence of a CNS depressant to screen for agents which may modulate the activity of the depressant on mGluR5 activity, if any.

After the identification of agents using cells in culture, the agents are tested in vivo for their effect on CNS depressant activity. Transgenic mice harboring the YRT2 sequences allow the agents to be tested in the context of the CNS and allow the use of assays for CNS depressant sensitivity as described above (e.g., LORR). The transgenic mice for this assay could be those as described in Example 3. As indicated in Example 3, the CNS depressant sensitivity of the transgenic mice appears to be of an intermediate level between that of SS mice and that of LS mice at which an agent can be tested for its ability to increase or decrease the CNS depressant sensitivity. Control over the heterozygosity or

the copy number of transgenic YRT2 sequences may provide an animal with a suitable level of CNS depressant sensitivity of this screen.

The agent is administered to animals with or without a CNS depressant (e.g., propofol). Determination of the CNS depressant activity of the agent is made using the standard LORR assessment protocols described in previous Examples.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity and understanding, it will be apparent to those skilled in the art that certain changes and modifications can be practiced. Therefore, the description and examples should not be construed as limiting the scope of the invention, which is delineated by the appended claims.

Table 3: Sleep times of transgenic receiving mGluR agonist and propofol

First Set of Transgenics Used in Study

Line	Cage	DOB	Mou	Color	weight	DOI for S	Inj PB	Inj CH	DOI for ST	Weight	Age	Δt^1	Agonist	T Inj Ag	Δt^2	Inj Prop (µl)	T out	T Up	ST	brain Vial	Hit ICV
Tg8xISS	1	7230	1M	Brown	20.1	7318	0.20	0.20	7322	19.4	92	4	5 µl	0.00	41.50	38.8	41.50	45.50	4.00	1	Y
Tg8xISS	2	7228	1M	brown	20.8	7318	0.21	0.21	7322	19.9	94	4	5 µl	51.00	42.50	39.8	93.50	96.00	2.50	3	Y
Tg8xISS	3	7250	2M	brown	22.3	7318	0.22	0.22	7322	22.2	74	4	5 µl	62.50	37.83	44.4	100.33	105.75	5.42	5	Y
																			Mean 3.97	SD 1.46	
Control	1	7230	3F	white	19.0	7318	0.19	0.19	7322	18.7	92	4	5 µl	12.00	?	37.4	NS	NS	0.00	2	Y
Control	2	7228	2F	white	18.3	7318	0.18	0.18	7322	20.8	94	4	5 µl	55.00	39.83	41.6	94.83	97.33	2.50	4	Y
Control	3	7250	1F	white	17.9	7318	0.18	0.18	7322	19.1	72	4	5 µl	59.50	?	38.2	NS	NS	0.00	CFO	X
																			Mean 0.83	SD 1.44	ΔMean 3.14
																					P-Value 0.028523

DOB = Date of Birth
DOI = Date of Injection
S = Surgery
PB = Pentobarbitol concentration of solution is 8 mg/ml
CH = Chloral Hydrate concentration of solution is 12 mg/ml
ST = Sleep Time, Time Up - Time Out (min)
Δt^1 = number of days in between surgery and propofol inj.
Agonist = ACPD 62.5 pmol/5µl
Prop = Propofol, concentration 20 mg/ml
Δt^2 = number of minutes in between agonist inj and prop inj
CFO = Cannula Fell out of Skull

Table 4.

LS and SS mice given Propofol and Agonist

	Line	Cage	DOB	Mou	weight	DOI for S	Inj PB	Inj CH	DOI for ST	Weight	Age	Δt^1	Agonist	T Inj Ag	Δt^2	T out/ Pr In	T Up	ST	brain Vial	Hit ICV	
LS F	LS 62-3A	1	7270	1F	20.5	7326	0.21	0.21	7329	19.2	59	3	5µl ACPD	0	33	33	39.75	6.75	1	Y	
ACPD	LS 62-5B	1	7270	2F	20.2	7326	0.2	0.2	7329	18.7	59	3	5µl ACPD	3.67	31.58	35.25	40.5	5.25	2	Y	
	LSP	2	7270	1F	17.7	7326	0.18	0.18	7329	17.5	59	3	5µl ACPD	6.25	31.33	37.58	41.58	4	4	Y	
	LSP	2	7270	3F	17.1	7326	0.17	0.17	7329	17.3	59	3	5µl ACPD	11	30.25	41.25	44.83	3.58	5	Y	
LS M ACPD	LSP	3	7270	1M	25.0	7326	0.25	0.25	7329	24.8	59	3	5µl ACPD	14.83	29.5	44.33	48.33	4	7	Y	
																	MEAN=	4.72			
																					P-Value= 0.239003
LS F	LS-62-5B	1	7270	3F	18.5	7326	0.19	0.19	7329	17.8	59	3	5µl saline	0	31.25	31.25	37.75	6.5	3	Y	
Saline	LSP	2	7270	4F	18.6	7326	0.19	0.19	7329	17.5	59	3	5µl saline	4.5	49.75	54.25	58.67	4.42	6	Y	
LS M Saline	LSP	3	7270	2M	23.6	7326	0.24	0.24	7329	23.4	59	3	5µl saline	8.33	34.42	42.75	48	5.25	6	Y	
																	MEAN=	5.39			
																					Δ Mean = 0.67
SS F	SS 62-1B	4	7268	2F	13.7	7326	0.14	0.14	7329	13.6	61	3	5µl ACPD	20	35.25	55.25	NS	0	10	Y	
ACPD	SS-62-4A	5	7268	1F	14.2	7326	0.14	0.14	7329	14.6	61	3	5µl ACPD	23.17	34.33	57.5	59.5	2	9	Y	
	SS-62-4A	5	7268	2F	15.3	7326	0.15	0.15	7329	15.5	61	3	5µl ACPD	26.33	35.34	61.67	63	1.33	12	Y	
																	MEAN=	1.11			
																					P-Value= 0.005142
SS F	SS 62-1B	4	7268	3F	16.3	7326	0.16	0.16	7329	16.4	61	3	5µl saline	11	35.25	46.25	50	3.75	11	Y	
SALINE	SS-62-4A	5	7268	3F	14.6	7326	0.15	0.15	7329	14.4	61	3	5µl saline	14.57	34.25	48.92	52.67	3.75	13	Y	
	SS-62-4A	5	7268	4F	17.0	7326	0.17	0.17	7329	17.7	61	3	5µl saline	18.83	33.25	52.08	56	3.92	14	Y	
																	MEAN=	3.81			
																					Δ Mean = 2.70

DOB = Date of Birth
DOI = Date of Injection
S = Surgery
PB = Pentobarbitol concentration of solution is 8 mg/ml (80 mg/kg).
CH = Chloral Hydrate concentration of solution is 12 mg/ml (120 mg/kg)
ST = Sleep Time, Time Up - Time Out (min)
Δt^1 = number of days in between surgery and propofol inj.
Agonist = ACPD 100 pmol/5µl
Prop = Propofol, concentration 20 mg/ml
Δt^2 = number of minutes in between agonist inj and prop inj
CFO = Cannula Fell out of Skull

CLAIMS

What is claimed is:

1. A method of identifying an agent which may exhibit CNS depressant activity, comprising the steps of:

(a) introducing a polynucleotide comprising a mouse polynucleotide sequence associated with CNS depressant sensitivity into a suitable host cell, wherein the mouse polynucleotide sequence corresponds to a polynucleotide sequence of yeast artificial chromosome YRT2;

(b) contacting host cell of step (a) with at least one agent to be tested;

(c) analyzing at least one characteristic associated with expression of the mouse YRT2 polynucleotide, wherein an agent is identified by its ability to modulate expression of the mouse YRT2 polynucleotide.

2. The method of claim 1, wherein the mouse polynucleotide is contained in an mGluR5 gene.

3. The method of claim 2, wherein the mouse polynucleotide is a control region of the mGluR5 gene.

4. A method of classifying a CNS depressant, comprising the steps of:

(a) introducing a polynucleotide comprising a mouse polynucleotide sequence associated with CNS depressant sensitivity into a suitable host cell, wherein the mouse polynucleotide sequence corresponds to a polynucleotide sequence of yeast artificial chromosome YRT2;

(b) contacting host cell of step (a) with at least one agent to be tested;

(c) analyzing at least one characteristic associated with expression of the mouse YRT2 polynucleotide, wherein an agent is identified by its ability to modulate expression of the mouse YRT2 polynucleotide, and wherein a CNS depressant which modulates expression may fall within a class of CNS depressants which displays differential reactivity in LS and SS mice.

5. The method of claim 4, wherein the mouse polynucleotide is contained in mGluR5 gene.

6. The method of claim 6, wherein the mouse polynucleotide is a control region of the mGluR5 gene.

7. A method of identifying an agent which may modulate CNS depressant sensitivity in an individual, said method comprising:

(a) introducing a polynucleotide comprising a mouse polynucleotide sequence associated with CNS depressant sensitivity into a suitable host cell, wherein the mouse polynucleotide sequence corresponds to a polynucleotide sequence of yeast artificial chromosome YRT2;

(b) contacting host cell of (a) with at least one agent to be tested;

(c) analyze at least one characteristic associated with expression of the mouse YRT2 polynucleotide, wherein an agent is identified by its ability to modulate expression of the mouse YRT2 polynucleotide.

8. The method of claim 7, wherein the mouse polynucleotide is contained in an mGluR5 gene.

9. The method of claim 8, wherein the mouse polynucleotide is a control region of the mGluR5 gene.

10. A method of identifying an agent which may exhibit CNS depressant activity, comprising the steps of:

(a) introducing a polynucleotide comprising a metabotropic glutamine receptor polynucleotide sequence into a suitable host cell;

(b) contacting host cell of step (a) with at least one agent to be tested:

(c) analyze at least one characteristic associated with expression of the mammalian metabotropic glutamate receptor polynucleotide, wherein an agent is identified by its ability to modulate expression of the mammalian metabotropic glutamate receptor polynucleotide.

11. The method of claim 10, wherein the metabotropic glutamate receptor polynucleotide is mammalian.

12. The method of claim 11, wherein the mammalian metabotropic glutamate receptor polynucleotide is human.

13. The method of claim 11, wherein the mammalian metabotropic glutamate receptor is mGluR5.

14. The method of claim 11, wherein the mammalian metabotropic glutamate receptor is mGluR1.

15. A method of classifying a CNS depressant, comprising the steps of:

(a) introducing a polynucleotide comprising a metabotropic glutamate receptor polynucleotide sequence into a suitable host cell;

(b) contacting host cell of step (a) with at least one agent to be tested;

(c) analyzing at least one characteristic associated with expression of the metabotropic glutamate receptor polynucleotide, wherein an agent is identified by its ability to modulate expression of the metabotropic glutamate receptor polynucleotide sequence, and wherein a CNS depressant which modulates the metabotropic glutamate receptor polynucleotide expression may fall within a class of CNS depressants which displays differential reactivity in LS and SS mice.

16. The method of claim 15, wherein the metabotropic glutamate receptor polynucleotide is mammalian.

17. The method of claim 16, wherein the mammalian metabotropic glutamate receptor polynucleotide is human.

18. The method of claim 17, wherein the human metabotropic glutamate receptor is mGluR5.

19. The method of claim 16, wherein the human metabotropic glutamate receptor is mGluR1.

20. An isolated polynucleotide comprising a polynucleotide encoding a mouse mGluR5 polypeptide, wherein the mouse mGluR5 polypeptide is at least about 10 contiguous amino acids of SEQ ID NO:2 and exhibits mGluR5 activity, and wherein the at least about 10 contiguous amino acids are not depicted within SEQ ID NO:6 or SEQ ID NO:7.

21. An isolated polynucleotide comprising a polynucleotide of at least about 25 contiguous nucleotides of SEQ ID NO:1, wherein the at least about 25 contiguous nucleotides are not depicted in SEQ NO: 3 or SEQ ID NO:5. --

22. An isolated polynucleotide comprising a region of at least about 25 contiguous nucleic acids of SEQ ID NO:1, said region having at least about 97% sequence identity to a sequence in SEQ ID NO:1.

23. A cloning vector comprising the polynucleotide of claim 20.

24. A cloning vector comprising the polynucleotide of claim 21.

25. A cloning vector comprising the polynucleotide of claim 22.

26. An expression vector comprising the polynucleotide of claim 20.

27. An expression vector comprising the polynucleotide of claim 21.

28. An expression vector comprising the polynucleotide of claim 22.

29. A host cell comprising the polynucleotide of claim 20.

30. A host cell comprising the polynucleotide of claim 21.

31. A host cell comprising the polynucleotide of claim 22.

32. The host cell of claim 29, wherein the host cell is mammalian.

33. The host cell of claim 30, wherein the host cell is mammalian.

34. The host cell of claim 31, wherein the host cell is mammalian.

35. An isolated polypeptide comprising at least about 5 contiguous amino acids of the sequence of SEQ ID NO:2. and wherein the at least about 5 contiguous amino aicds are not depicted in SEQ ID NO:6 or SEQ ID NO:7.

36. The polypeptide of claim 35, wherein the polypeptide comprises the sequence of SEQ ID NO:2.

37. A purified antibody capable of specifically binding to a polypeptide of claim 35.

38. A monoclonal antibody capable of specifically binding to a polypeptide of claim 35.

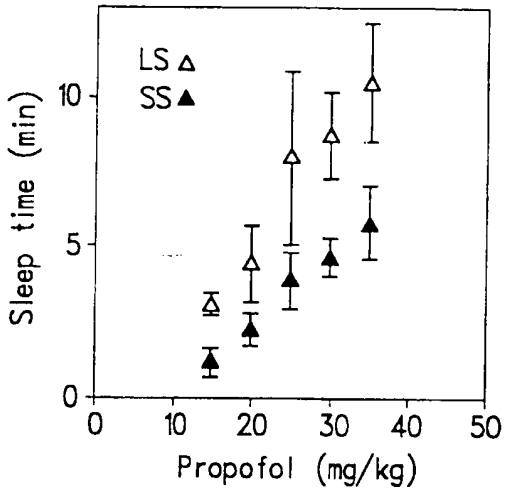
LS △
SS ▲
Sleep time (min)
0
5
10
0
10
20
30
40
50
Propofol (mg/kg)

FIG. 1

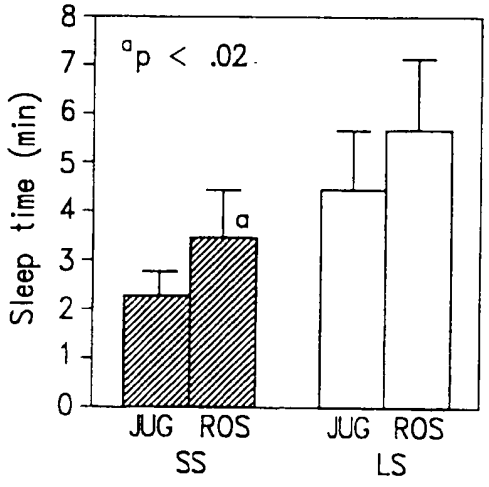
$^{a}p < .02$
Sleep time (min)
0
1
2
3
4
5
6
7
8
a
JUG ROS
SS
JUG ROS
LS

FIG. 2

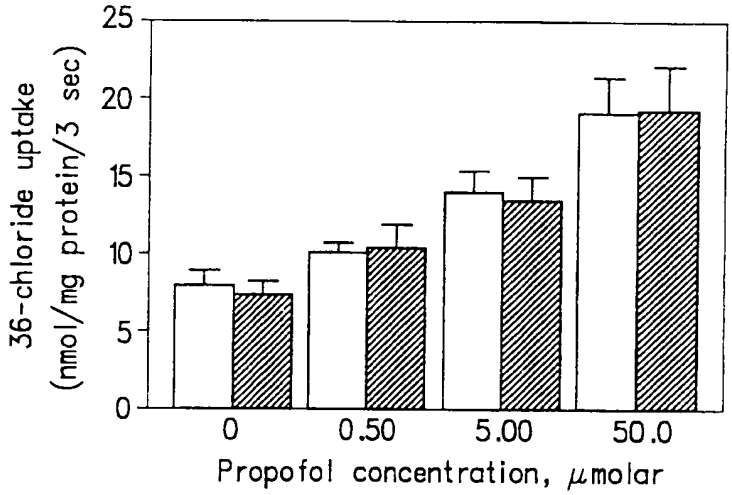
36-chloride uptake
(nmol/mg protein/3 sec)
25
20
15
10
5
0
0
0.50
5.00
50.0
Propofol concentration, μmolar

FIG. 3

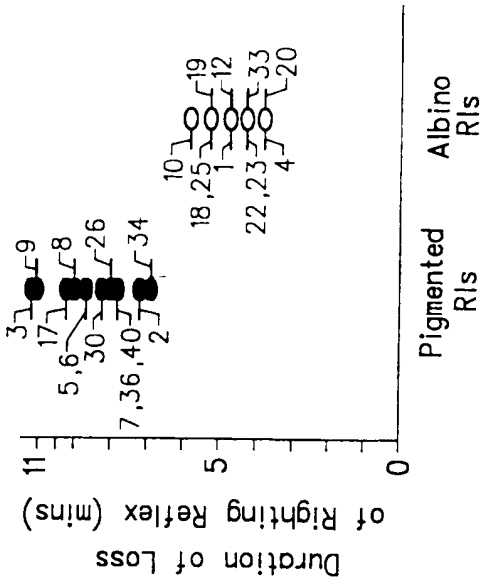
Duration of Loss
of Righting Reflex (mins)
11
5
0
3
17
5,6
30
7,36,40
2
9
8
26
34
10
18,25
1
22,23
4
19
12
33
20
Pigmented Rls
Albino Rls

FIG. 4B

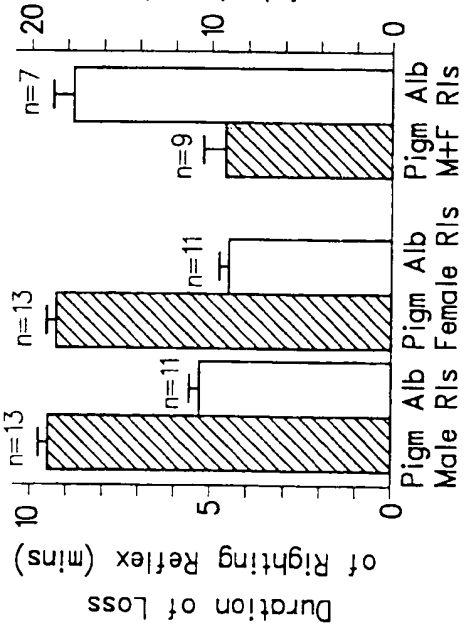
Duration of Loss
of Righting Reflex (mins)
10
5
0
n=13
n=11
n=13
n=11
Pigm Alb Pigm Alb
Male Rls Female Rls
Propofol at awakening
(μg/gm brain)
20
10
0
n=7
n=9
Pigm Alb
M+F Rls

FIG. 4A

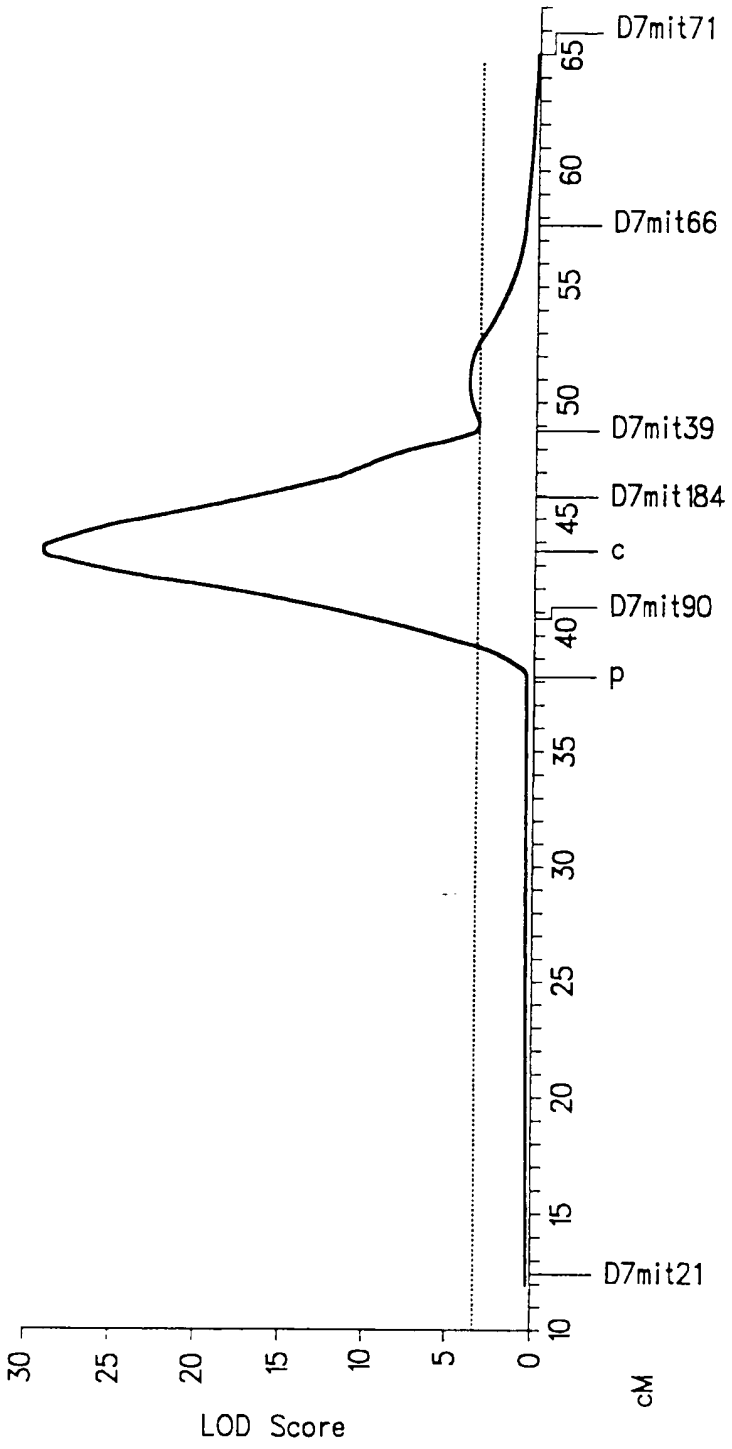
D7mit71
D7mit66
D7mit39
D7mit184
c
D7mit90
p
D7mit21
65
60
55
50
45
40
35
30
25
20
15
10
0
5
10
15
20
25
30
cM
LOD Score

FIG. 5

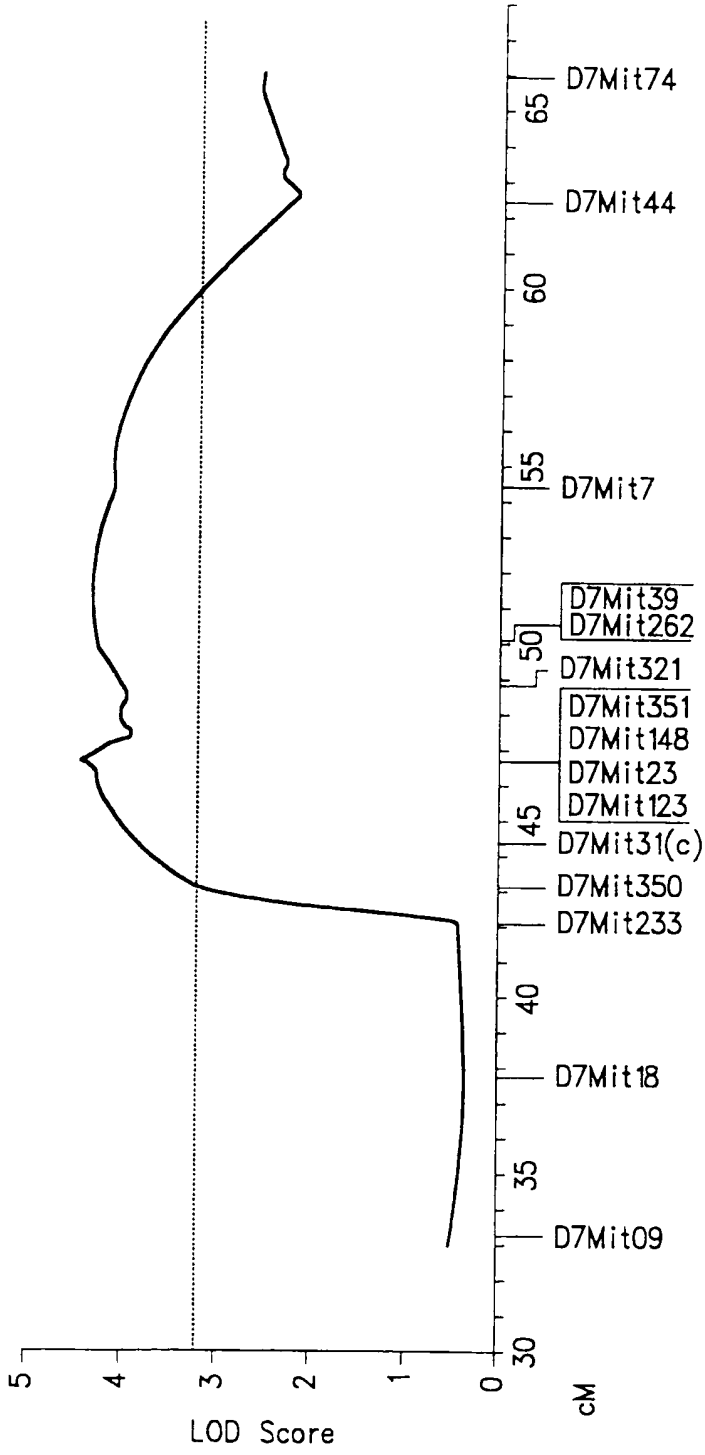
D7Mit74
D7Mit44
D7Mit7
D7Mit39
D7Mit262
D7Mit321
D7Mit351
D7Mit148
D7Mit23
D7Mit123
D7Mit31(c)
D7Mit350
D7Mit233
D7Mit18
D7Mit09
65
60
55
50
45
40
35
30
cM
0
1
2
3
4
5
LOD Score

FIG. 6

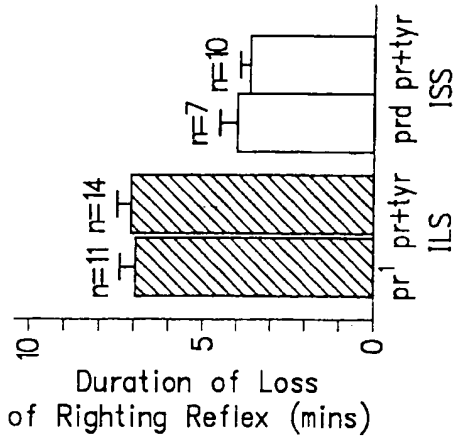
n=11 n=14
n=7 n=10
10
5
0
Duration of Loss
of Righting Reflex (mins)
pr1 pr+tyr prd pr+tyr
ILS
ISS

FIG. 7B

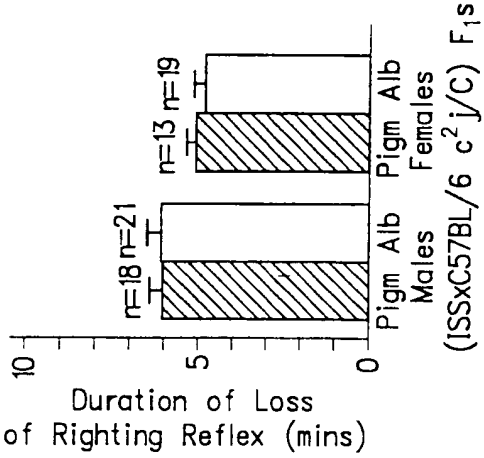
n=18 n=21
n=13 n=19
10
5
0
Duration of Loss
of Righting Reflex (mins)
Pigm Alb
Males
Pigm Alb
Females
(ISSxC57BL/6 c2j/C) F1s

FIG. 7A

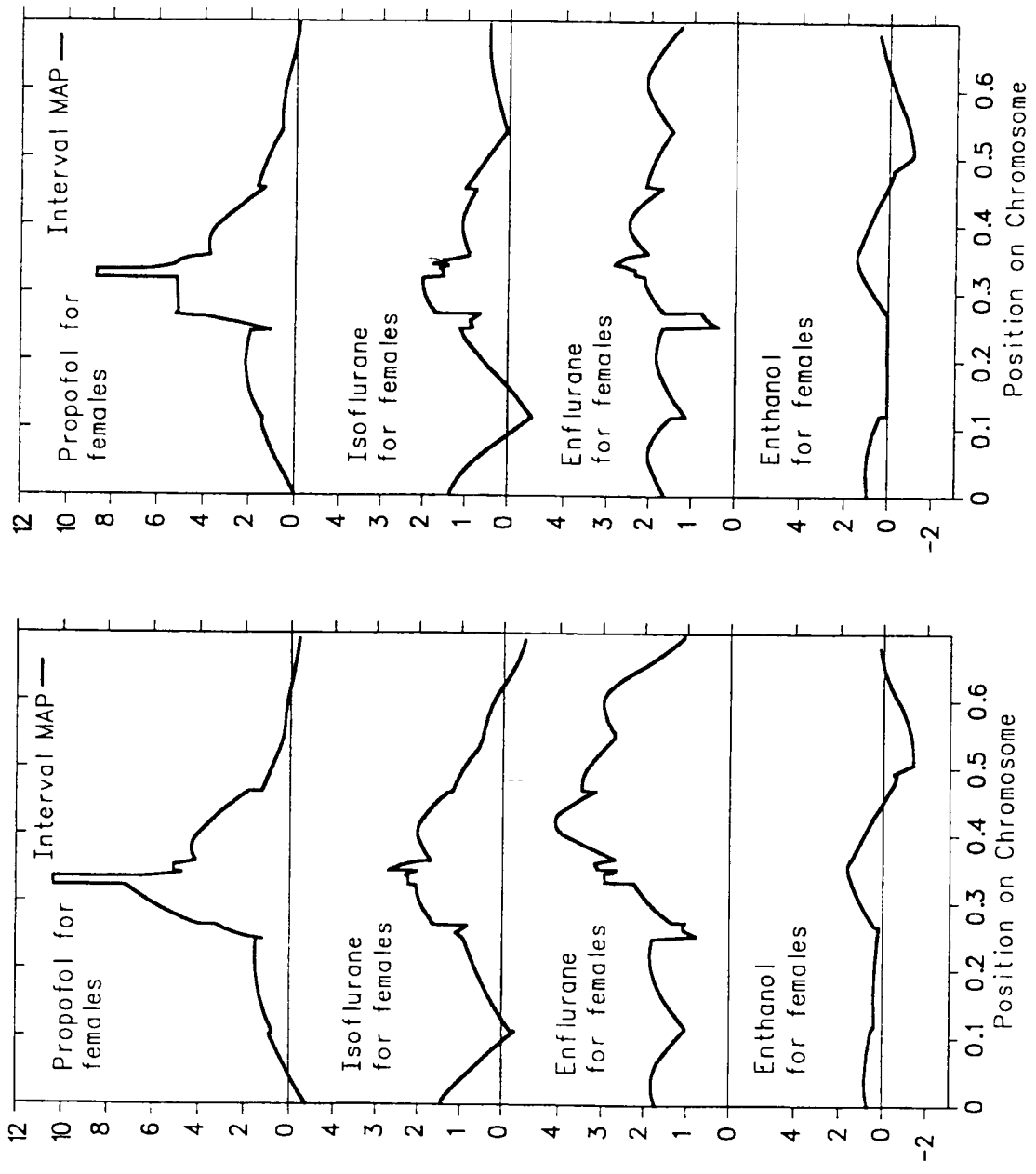
Propofol for females
Interval MAP
Isoflurane for females
Enflurane for females
Enthanol for females
Position on Chromosome

FIG. 8

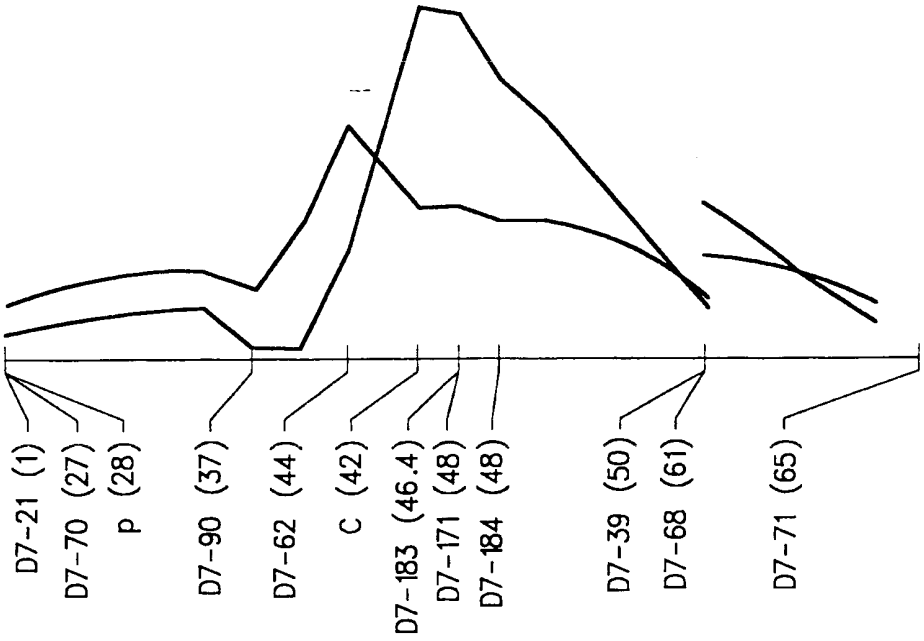
D7-21 (1)
D7-70 (27)
P (28)
D7-90 (37)
D7-62 (44)
C (42)
D7-183 (46.4)
D7-171 (48)
D7-184 (48)
D7-39 (50)
D7-68 (61)
D7-71 (65)

FIG. 9

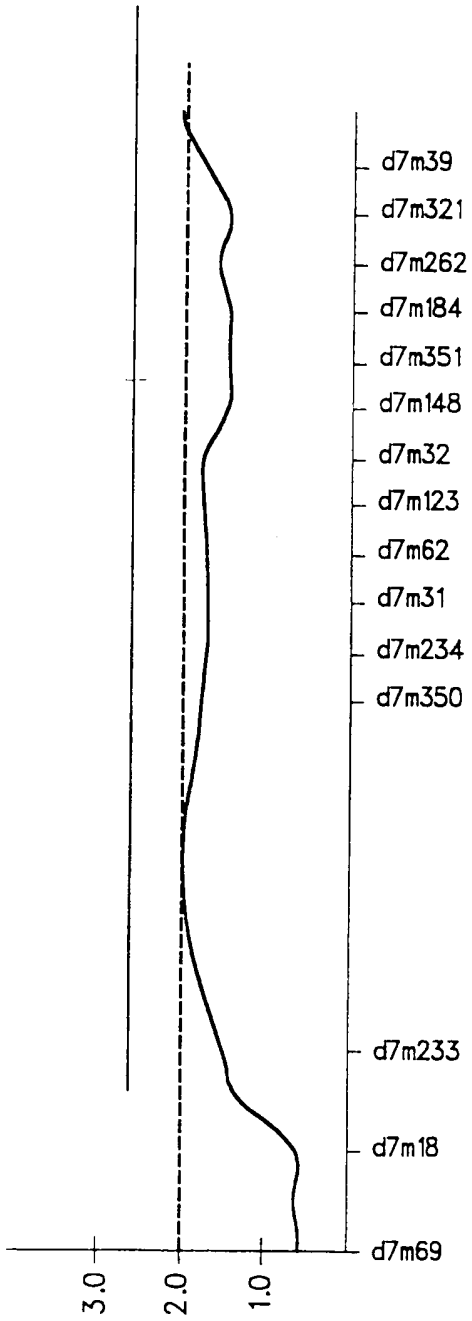
d7m39
d7m321
d7m262
d7m184
d7m351
d7m148
d7m32
d7m123
d7m62
d7m31
d7m234
d7m350
d7m233
d7m18
d7m69
3.0
2.0
1.0

FIG. 10

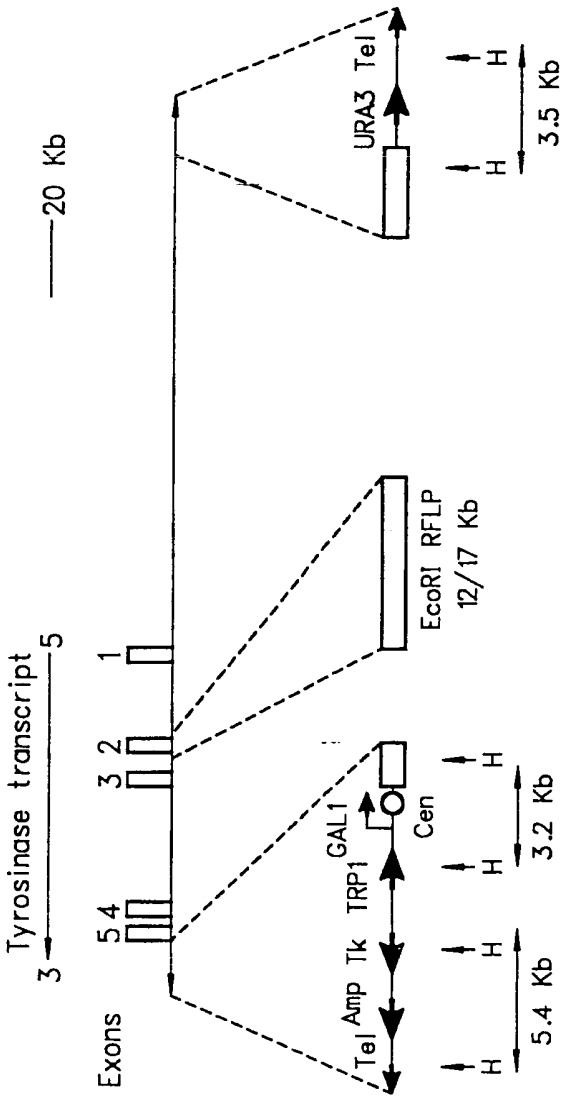
20 Kb
Tyrosinase transcript
3
5
Exons
5 4
3 2
1
Tel
Amp
Tk
TRP1
GAL1
Cen
URA3
Tel
EcoRI RFLP
12/17 Kb
H
H
5.4 Kb
H
H
3.2 Kb
H
H
3.5 Kb

FIG. 11

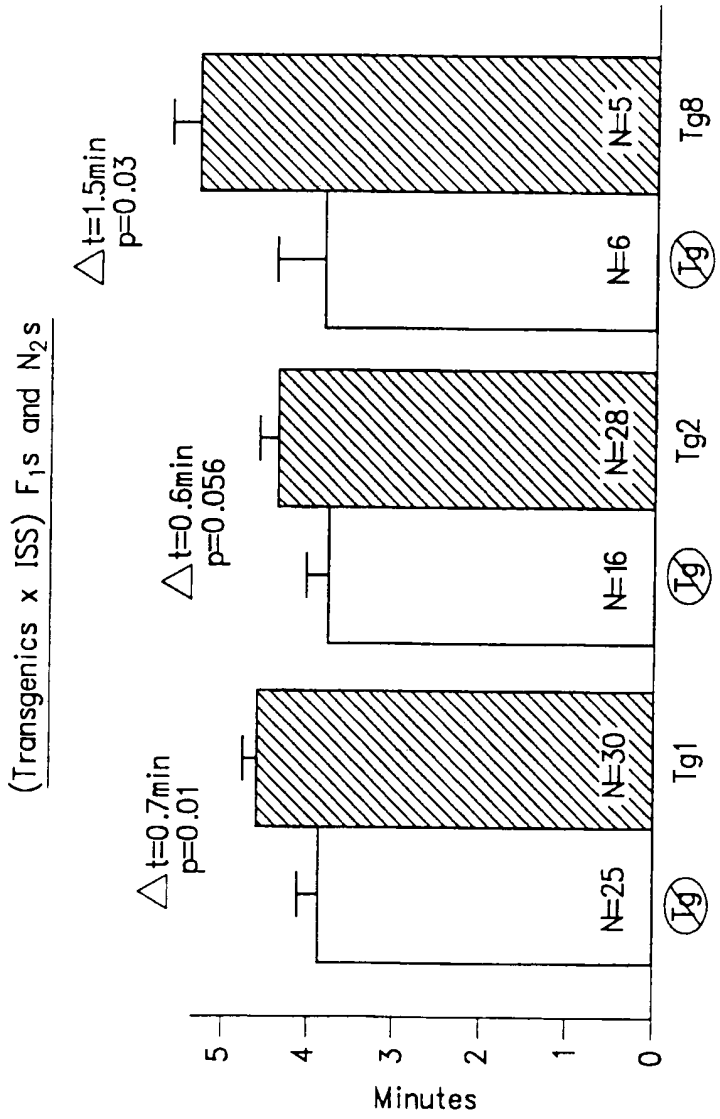
(Transgenics x ISS) F_1s and N_2s
Δt=0.7min
p=0.01
Δt=0.6min
p=0.056
Δt=1.5min
p=0.03
N=25
N=30
N=16
N=28
N=6
N=5
Tg
Tg1
Tg2
Tg8
0
1
2
3
4
5
Minutes

FIG. 12

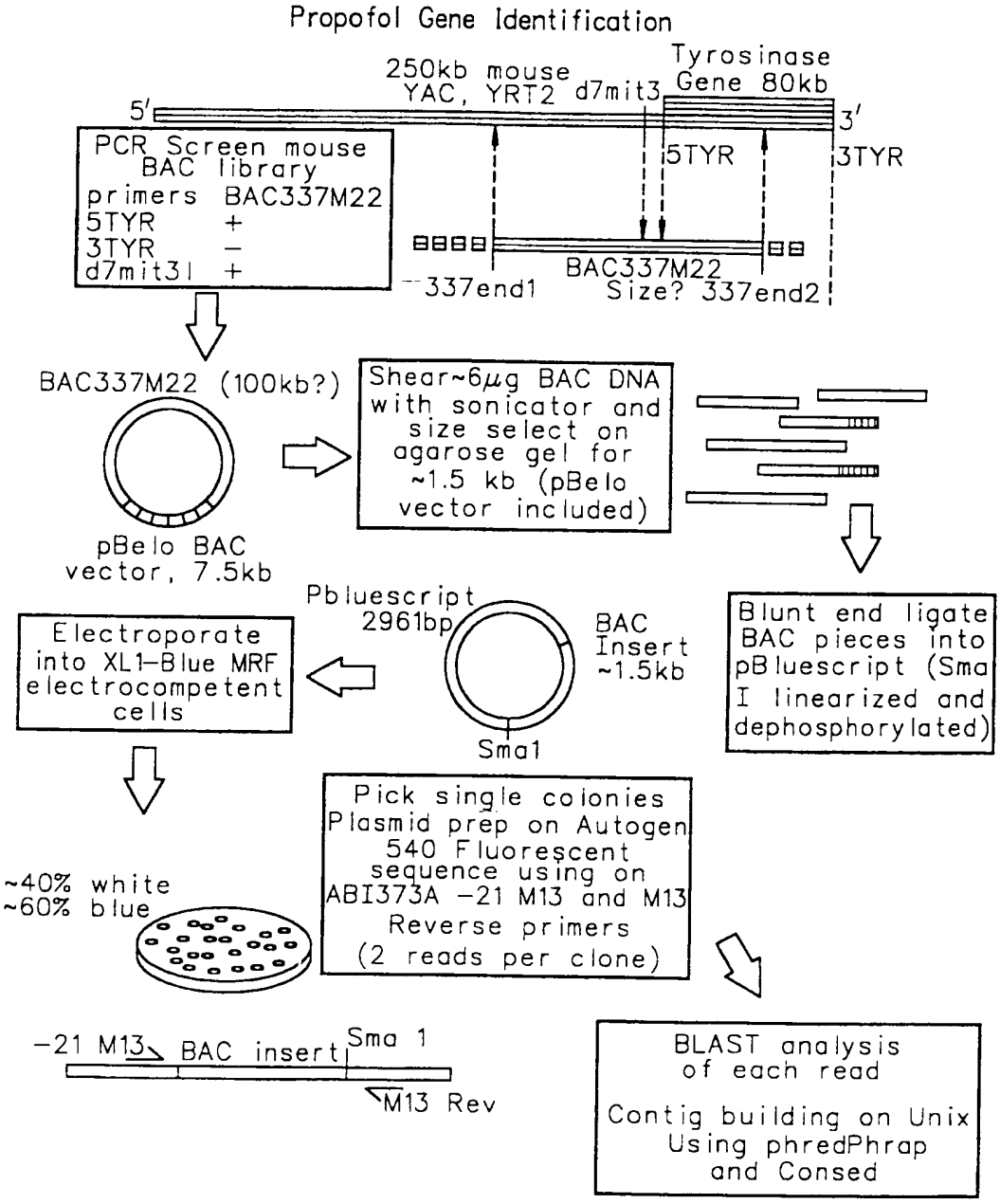
Propofol Gene Identification
250kb mouse YAC, YRT2 d7mit3
Tyrosinase Gene 80kb
5'
3'
5TYR
3TYR
PCR Screen mouse BAC library
primers BAC337M22
5TYR +
3TYR –
d7mit31 +
337end1
BAC337M22 Size? 337end2
BAC337M22 (100kb?)
pBelo BAC vector, 7.5kb
Shear~6μg BAC DNA with sonicator and size select on agarose gel for ~1.5 kb (pBelo vector included)
Blunt end ligate BAC pieces into pBluescript (Sma I linearized and dephosphorylated)
Pbluescript 2961bp
BAC Insert ~1.5kb
Sma1
Electroporate into XL1-Blue MRF electrocompetent cells
~40% white
~60% blue
Pick single colonies Plasmid prep on Autogen 540 Fluorescent sequence using on ABI373A –21 M13 and M13 Reverse primers (2 reads per clone)
–21 M13
BAC insert
Sma 1
M13 Rev
BLAST analysis of each read
Contig building on Unix Using phredPhrap and Consed

FIG. 13

mouse mGluR5

```
 ATG GTC CTT CTG TTG ATT CTG TCA GTC CTA CTT CTG AAA GAA GAT GTA
48
 Met Val Leu Leu Leu Ile Leu Ser Val Leu Leu Leu Lys Glu Asp Val
  1               5                  10                  15

 CGA GGA AGT GCA CAG TCC AGT GAG AGG AGG GTG GTG GCT CAC ATG CCA
96
 Arg Gly Ser Ala Gln Ser Ser Glu Arg Arg Val Val Ala His Met Pro
             20                  25                  30

 GGT GAC ATT ATT ATT GGA GCT CTC TTC TCG GTC CAC CAC CAA CCA ACT
144
 Gly Asp Ile Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro Thr
         35                  40                  45

 GTG GAC AAA GTT CAT GAG AGG AAG TGT GGG GCA GTC CGT GAG CAG TAT
192
 Val Asp Lys Val His Glu Arg Lys Cys Gly Ala Val Arg Glu Gln Tyr
     50                  55                  60

 GGC ATT CAG AGA GTG GAA GCC ATG CTG CAT ACC TTG GAA AGG ATC AAT
240
 Gly Ile Gln Arg Val Glu Ala Met Leu His Thr Leu Glu Arg Ile Asn
 65                  70                  75                  80

 TCA GAT CCC ACA CTC TTG CCC AAC ATC ACA CTA GGC TGT GAG ATA AGA
288
 Ser Asp Pro Thr Leu Leu Pro Asn Ile Thr Leu Gly Cys Glu Ile Arg
                 85                  90                  95

 GAT TCC TGC TGG CAT TCT GCT GTG GCC CTA GAG CAA AGC ATT GAG TTT
336
 Asp Ser Cys Trp His Ser Ala Val Ala Leu Glu Gln Ser Ile Glu Phe
             100                 105                 110

 ATA AGG GAT TCC CTC ATC TCT TCG GAA GAG GAA GAA GGC TTG GTA CGC
384
 Ile Arg Asp Ser Leu Ile Ser Ser Glu Glu Glu Glu Gly Leu Val Arg
         115                 120                 125

 TGT GTA GAT GGC TCT TCT TCC TTC CGC TCC AAG AAA CCC ATA GTG GGA
432
 Cys Val Asp Gly Ser Ser Ser Phe Arg Ser Lys Lys Pro Ile Val Gly
     130                 135                 140

 GTC ATT GGG CCT GGC TCG AGT TCT GTG GCC ATT CAA GTT CAG AAC TTG
480
 Val Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val Gln Asn Leu
 145                 150                 155                 160

 CTC CAG CTT TTC AAC ATA CCT CAG ATT GCT TAC TCT GCA ACT AGC ATG
528
 Leu Gln Leu Phe Asn Ile Pro Gln Ile Ala Tyr Ser Ala Thr Ser Met
                 165                 170                 175
```

FIGURE 14(A)

```
 GAT TTG AGT GAC AAG ACT CTA TTC AAG TAC TTC ATG AGG GTT GTA CCT
576
 Asp Leu Ser Asp Lys Thr Leu Phe Lys Tyr Phe Met Arg Val Val Pro
             180                 185                 190

 TCG GAT GCC CAG CAA GCC CGA GCC ATG GTA GAC ATA GTG AAG AGA TAC
624
 Ser Asp Ala Gln Gln Ala Arg Ala Met Val Asp Ile Val Lys Arg Tyr
         195                 200                 205

 AAC TGG ACT TAT GTC TCA GCT GTG CAC ACA GAA GGC AAC TAT GGA GAA
672
 Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn Tyr Gly Glu
     210                 215                 220

 AGT GGG ATG GAG GCT TTC AAA GAT ATG TCA GCG AAG GAA GGG ATT TGC
720
 Ser Gly Met Glu Ala Phe Lys Asp Met Ser Ala Lys Glu Gly Ile Cys
 225                 230                 235                 240

 ATC GCC CAC TCT TAC AAA ATC TAC AGC AAT GCT GGG GAA CAG AGC TTT
768
 Ile Ala His Ser Tyr Lys Ile Tyr Ser Asn Ala Gly Glu Gln Ser Phe
                 245                 250                 255

 GAC AAG CTG TTG AAA AAG CTC AGA AGT CAT TTA CCT AAA GCC CGG GTG
816
 Asp Lys Leu Leu Lys Lys Leu Arg Ser His Leu Pro Lys Ala Arg Val
             260                 265                 270

 GTA GCC TGC TTC TGT GAA GGC ATG ACA GTT CGA GGT CTG CTC ATG GCC
864
 Val Ala Cys Phe Cys Glu Gly Met Thr Val Arg Gly Leu Leu Met Ala
         275                 280                 285

 ATG AGA CGC TTG GGT CTA GCA GGG GAA TTT CTA CTT CTG GGC AGT GAT
912
 Met Arg Arg Leu Gly Leu Ala Gly Glu Phe Leu Leu Leu Gly Ser Asp
     290                 295                 300

 GGC TGG GCT GAC AGG TAT GAC GTG ACA GAT GGG TAT CAG CGA GAA GCT
960
 Gly Trp Ala Asp Arg Tyr Asp Val Thr Asp Gly Tyr Gln Arg Glu Ala
 305                 310                 315                 320

 GTC GGT GGG ATT ACA ATC AAG CTC CAG TCT CCT GAT GTC AAG TGG TTT
1008
 Val Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro Asp Val Lys Trp Phe
                 325                 330                 335

 GAT GAT TAT TAT CTG AAG CTC CGG CCA GAA ACA AAC CTC AGA AAC CCT
1056
 Asp Asp Tyr Tyr Leu Lys Leu Arg Pro Glu Thr Asn Leu Arg Asn Pro
             340                 345                 350
```

FIGURE 14(B)

```
 TGG TTT CAA GAA TTT TGG CAG CAT CGT TTT CAG TGC CGG CTA GAA GGG
1104
 Trp Phe Gln Glu Phe Trp Gln His Arg Phe Gln Cys Arg Leu Glu Gly
         355                 360                 365

 TTT GCA CAG GAG AAC AGC AAG TAC AAC AAG ACT TGC AAC AGT TCT CTA
1152
 Phe Ala Gln Glu Asn Ser Lys Tyr Asn Lys Thr Cys Asn Ser Ser Leu
     370                 375                 380

 ACT CTG AGA ACG CAT CAT GTT CAA GAT TCC AAA ATG GGA TTT GTG ATC
1200
 Thr Leu Arg Thr His His Val Gln Asp Ser Lys Met Gly Phe Val Ile
 385                 390                 395                 400

 AAT GCA ATC TAT TCT ATG GCT TAT GGG CTC CAC AAC ATG CAG ATG TCC
1248
 Asn Ala Ile Tyr Ser Met Ala Tyr Gly Leu His Asn Met Gln Met Ser
                 405                 410                 415

 CTG TGT CCA GGC TAT GCA GGC CTC TGT GAT GCA ATG AAG CCA ATT GAT
1296
 Leu Cys Pro Gly Tyr Ala Gly Leu Cys Asp Ala Met Lys Pro Ile Asp
             420                 425                 430

 GGG CGG AAA CTT TTG GAC TCC CTG ATG AAA ACC AAC TTT ACT GGA GTT
1344
 Gly Arg Lys Leu Leu Asp Ser Leu Met Lys Thr Asn Phe Thr Gly Val
         435                 440                 445

 TCC GGA GAT ATG ATT CTA TTT GAT GAA AAT GGA GAC TCT CCA GGA AGG
1392
 Ser Gly Asp Met Ile Leu Phe Asp Glu Asn Gly Asp Ser Pro Gly Arg
     450                 455                 460

 TAT GAA ATA ATG AAT TTC AAG GAA ATG GGA AAA GAT TAT TTT GAT TAC
1440
 Tyr Glu Ile Met Asn Phe Lys Glu Met Gly Lys Asp Tyr Phe Asp Tyr
 465                 470                 475                 480

 ATC AAT GTT GGA AGT TGG GAC AAT GGG GAA TTA AAA ATG GAT GAT GAT
1488
 Ile Asn Val Gly Ser Trp Asp Asn Gly Glu Leu Lys Met Asp Asp Asp
                 485                 490                 495

 GAA GTG TGG TCC AAG AAA AAT AAC ATC ATC AGA TCT GTG TGC AGT GAA
1536
 Glu Val Trp Ser Lys Lys Asn Asn Ile Ile Arg Ser Val Cys Ser Glu
             500                 505                 510

 CCG TGT GAG AAA GGA CAG ATA AAG GTG ATC CGG AAG GGA GAA GTC AGC
1584
 Pro Cys Glu Lys Gly Gln Ile Lys Val Ile Arg Lys Gly Glu Val Ser
         515                 520                 525
```

FIGURE 14(C)

```
 TGT TGT TGG ACC TGC ACA CCT TGT AAG GAG AAT GAG TAT GTT TTT GAT
1632
 Cys Cys Trp Thr Cys Thr Pro Cys Lys Glu Asn Glu Tyr Val Phe Asp
     530                 535                 540

 GAG TAC ACC TGC AAG GCG TGC CAA CTG GGG TCC TGG CCC ACT GAC GAC
1680
 Glu Tyr Thr Cys Lys Ala Cys Gln Leu Gly Ser Trp Pro Thr Asp Asp
 545                 550                 555                 560

 TTG ACA GGT TGT GAT TTG ATC CCA GTC CAG TAT CTT CGA TGG GGT GAC
1728
 Leu Thr Gly Cys Asp Leu Ile Pro Val Gln Tyr Leu Arg Trp Gly Asp
                 565                 570                 575

 CCT GAG CCC ATT GCA GCT GTG GTG TTT GCC TGC CTC GGT CTG CTA GCC
1776
 Pro Glu Pro Ile Ala Ala Val Val Phe Ala Cys Leu Gly Leu Leu Ala
             580                 585                 590

 ACC CTC TTC GTT ACT GTA ATC TTC ATC ATT TAT CGG GAC ACT CCA GTG
1824
 Thr Leu Phe Val Thr Val Ile Phe Ile Ile Tyr Arg Asp Thr Pro Val
         595                 600                 605

 GTC AAG TCT TCC AGC AGG GAA CTC TGC TAC ATT ATC CTT GCT GGC ATC
1872
 Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile Ile Leu Ala Gly Ile
     610                 615                 620

 TGC CTG GGT TAC TTG TGT ACC TTC TGC CTC ATT GCA AAG CCC AAA CAG
1920
 Cys Leu Gly Tyr Leu Cys Thr Phe Cys Leu Ile Ala Lys Pro Lys Gln
 625                 630                 635                 640

 ATT TAC TGC TAC CTT CAG AGA ATT GGC ATC GGT CTC TCT CCA GCC ATG
1968
 Ile Tyr Cys Tyr Leu Gln Arg Ile Gly Ile Gly Leu Ser Pro Ala Met
                 645                 650                 655

 AGC TAC TCA GCC CTT GTA ACG AAG ACC AAC CGT ATT GCA AGG ATT CTA
2016
 Ser Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala Arg Ile Leu
             660                 665                 670

 GCT GGC AGC AAG AAG AAG ATC TGT ACC AAA AAG CCC AGA TTC ATG AGC
2064
 Ala Gly Ser Lys Lys Lys Ile Cys Thr Lys Lys Pro Arg Phe Met Ser
         675                 680                 685

 GCC TGT GCT CAG TTA GTG ATC GCT TTC ATT CTC ATC TGT ATA CAG TTG
2112
 Ala Cys Ala Gln Leu Val Ile Ala Phe Ile Leu Ile Cys Ile Gln Leu
     690                 695                 700
```

FIGURE 14(D)

```
 GGT ATT ATT GTG GCC CTC TTT ATC ATG GAG CCT CCG GAT ATA ATG CAT
2160
 Gly Ile Ile Val Ala Leu Phe Ile Met Glu Pro Pro Asp Ile Met His
 705                 710                 715                 720

 GAC TAT CCA AGC ATC CGA GAA GTC TAC TTG ATT TGT AAC ACC ACC AAC
2208
 Asp Tyr Pro Ser Ile Arg Glu Val Tyr Leu Ile Cys Asn Thr Thr Asn
                 725                 730                 735

 CTA GGG GTT GTC ACT CCT CTT GGA TAC AAT GGA TTA TTG ATT TTG AGT
2256
 Leu Gly Val Val Thr Pro Leu Gly Tyr Asn Gly Leu Leu Ile Leu Ser
             740                 745                 750

 TGC ACA TTC TAT GCG TTC AAG ACC AGA AAT GTT CCA GCC AAC TTT AAC
2304
 Cys Thr Phe Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala Asn Phe Asn
         755                 760                 765

 GAG GCC AAA TAT ATT GCT TTC ACC ATG TAC ACA ACC TGC ATC ATA TGG
2352
 Glu Ala Lys Tyr Ile Ala Phe Thr Met Tyr Thr Thr Cys Ile Ile Trp
     770                 775                 780

 CTG GCC TTT GTG CCT ATC TAC TTT GGC AGC AAC TAC AAA ATC ATC ACC
2400
 Leu Ala Phe Val Pro Ile Tyr Phe Gly Ser Asn Tyr Lys Ile Ile Thr
 785                 790                 795                 800

 ATG TGT TTC TCA GTC AGC CTC AGT GCC ACA GTG GCC CTG GGT TGC ATG
2448
 Met Cys Phe Ser Val Ser Leu Ser Ala Thr Val Ala Leu Gly Cys Met
                 805                 810                 815

 TTT GTG CCG AAG GTG TAC ATC ATC CTA GCC AAA CCG GAG AGA AAT GTG
2496
 Phe Val Pro Lys Val Tyr Ile Ile Leu Ala Lys Pro Glu Arg Asn Val
             820                 825                 830

 CGC AGC GCC TTC ACA ACC TCT ACA GTG GTG CGC ATG CAC GTA GGA GAT
2544
 Arg Ser Ala Phe Thr Thr Ser Thr Val Val Arg Met His Val Gly Asp
         835                 840                 845

 GGC AAG TCA TCA TCC GCT GCC AGC AGA TCC AGC AGC CTA GTC AAC CTG
2592
 Gly Lys Ser Ser Ser Ala Ala Ser Arg Ser Ser Ser Leu Val Asn Leu
     850                 855                 860

 TGG AAG AGG AGG GGC TCG TCT GGG GAA ACC CTA AGG TAC AAA GAC AGG
2640
 Trp Lys Arg Arg Gly Ser Ser Gly Glu Thr Leu Arg Tyr Lys Asp Arg
 865                 870                 875                 880
```

FIGURE 14(E)

```
 AGA CTG GCC CAG CAC AAG TCG GAA ATA GAG TGT TTC ACC CCC AAA GGG
2688
 Arg Leu Ala Gln His Lys Ser Glu Ile Glu Cys Phe Thr Pro Lys Gly
                 885                 890                 895

 AGT ATG GGG AAT GGT GGG AGA GCA ACA ATG AGC AGC TCC AAT GGA AAA
2736
 Ser Met Gly Asn Gly Gly Arg Ala Thr Met Ser Ser Ser Asn Gly Lys
             900                 905                 910

 TCC GTG ACT TGG GCC CAG AAT GAG AAG AGC ACC CGG GGC CAG CAC CTG
2784
 Ser Val Thr Trp Ala Gln Asn Glu Lys Ser Thr Arg Gly Gln His Leu
         915                 920                 925

 TGG CAG CGA CTG TCT GTC CAT ATC AAC AAG AAG GAG AAC CCC AAC CAG
2832
 Trp Gln Arg Leu Ser Val His Ile Asn Lys Lys Glu Asn Pro Asn Gln
     930                 935                 940

 ACA GCA GTC ATC AAA CCC TTC CCC AAG AGC ACA GAG AGC CGC GGG CAG
2880
 Thr Ala Val Ile Lys Pro Phe Pro Lys Ser Thr Glu Ser Arg Gly Gln
 945                 950                 955                 960

 GGT GCA GGG GCA GGT GGT GGC TCT GGC CCC GGT GCA GCT GGT GCT GGT
2928
 Gly Ala Gly Ala Gly Gly Gly Ser Gly Pro Gly Ala Ala Gly Ala Gly
                 965                 970                 975

 AGC GCA GGA TGC ACA GCG ACA GGC GGC CCA GAG CCA CCA GAC GCC GGC
2976
 Ser Ala Gly Cys Thr Ala Thr Gly Gly Pro Glu Pro Pro Asp Ala Gly
             980                 985                 990

 CCC AAG GCG CTT TAT GAT GTC GCA GAG GCA GAG GAG CGC TTC CCA GCG
3024
 Pro Lys Ala Leu Tyr Asp Val Ala Glu Ala Glu Glu Arg Phe Pro Ala
         995                1000                1005

 GCT GCC AGG CCG CGC TCG CCA TCG CCC ATC AGT ACG CTG AGC CAC CTG
3072
 Ala Ala Arg Pro Arg Ser Pro Ser Pro Ile Ser Thr Leu Ser His Leu
    1010                1015                1020

 GCA GGC TCG GCG GGC CGC ACA GAC GAC GAC GCG CCG TCG CTG CAC TCG
3120
 Ala Gly Ser Ala Gly Arg Thr Asp Asp Asp Ala Pro Ser Leu His Ser
1025                1030                1035                1040

 GAG ACC GCT GCA CGC AGC AGC TCA TCC CAG GGC TCG CTC ATG GAG CAG
3168
 Glu Thr Ala Ala Arg Ser Ser Ser Ser Gln Gly Ser Leu Met Glu Gln
                1045                1050                1055
```

FIGURE 14(F)

```
 ATT AGC AGC GTG GTG ACG CGC TTC ACC GCC AAC ATC ACC GAG CTC AAC
3216
 Ile Ser Ser Val Val Thr Arg Phe Thr Ala Asn Ile Thr Glu Leu Asn
           1060                1065                1070

 TCC ATG ATG TTG TCC ACC GCG GCT GCG CCG GGG CCC CCT GGT ACC CCT
3264
 Ser Met Met Leu Ser Thr Ala Ala Ala Pro Gly Pro Pro Gly Thr Pro
       1075                1080                1085

 ATC TGC TCT TCC TAC CTG ATC CCC AAA GAG ATC CAG CTG CCC ACG ACC
3312
 Ile Cys Ser Ser Tyr Leu Ile Pro Lys Glu Ile Gln Leu Pro Thr Thr
   1090                1095                1100

 ATG ACG ACC TTC GCA GAG ATC CAG CCT CTG CCG GCC ATC GAG GTG ACC
3360
 Met Thr Thr Phe Ala Glu Ile Gln Pro Leu Pro Ala Ile Glu Val Thr
1105                1110                1115                1120

 GGA GGA GCT CAG CCG GCG ACA GGG CCA TCA CCT GCC CAA GAG ACG CCC
3408
 Gly Gly Ala Gln Pro Ala Thr Gly Pro Ser Pro Ala Gln Glu Thr Pro
               1125                1130                1135

 GCA GGA GCT GAA GCC GCC CCA GGA AAA CCG GAT CTG GAG GAG CTG GTG
3456
 Ala Gly Ala Glu Ala Ala Pro Gly Lys Pro Asp Leu Glu Glu Leu Val
           1140                1145                1150

 GCC CTC ACT CCA CCA TCG CCC TTC AGG GAC TCG GTG GAC TCG GGG AGC
3504
 Ala Leu Thr Pro Pro Ser Pro Phe Arg Asp Ser Val Asp Ser Gly Ser
       1155                1160                1165

 ACC ACC CCA AAC TCT CCA GTC TCC GAA TCG GCC CTC TGC ATC CCA TCC
3552
 Thr Thr Pro Asn Ser Pro Val Ser Glu Ser Ala Leu Cys Ile Pro Ser
   1170                1175                1180

 TCT CCC AAA TAT GAC ACT CTC ATC ATC AGA GAT TAC ACG
3591
 Ser Pro Lys Tyr Asp Thr Leu Ile Ile Arg Asp Tyr Thr
1185                1190                1195
```

FIGURE 14(G)

Portion of mouse mGluR5 from Genebank database

```
 AG GTA CAA AGA CAG GAG ACT GGC CCA GCA CAA GTC GGA AAT AGA GTG
47
    Val Gln Arg Gln Glu Thr Gly Pro Ala Gln Val Gly Asn Arg Val
     1               5                  10                  15

 TTT CAC CCC CAA AGG GAG TAT GGG GAA TGG TGG GAG AGC AAC AAT GAG
95
 Phe His Pro Gln Arg Glu Tyr Gly Glu Trp Trp Glu Ser Asn Asn Glu
                 20                  25                  30

 CAG GT
100
 Gln
```

FIGURE 15

rat mGluR5

```
 TCAGACATCT CTTGAGAAAC TCAGCCTAGT TTATCCAGCT GTTTTGTCCA CATATGTTCA
60
 TAAAATCTAG TTATGAATAT TCTGATCGTA TGGATGCACC TAACTCGACA GTTTCTGGGG
120
 ACTTTGCTAT CTTTTAGAGA ATATATCTAA ATTGCTGACT AATTCCTCAA TGCGCCACTC
180
 AACGCAGGGC ATCATTTGGC TGTACCTATT GGAGCTCTGA GTTTCCTCCA TCATGATATC
240
 TTTATTAGCT TGAATTCCTT TCCTAAA ATG GTC CTT CTG TTG ATC CTG TCA GTC
294
                              Met Val Leu Leu Leu Ile Leu Ser Val
                               1               5

 CTA CTT CTG AAA GAA GAT GTA CGA GGG AGT GCA CAG TCC AGT GAG AGG
342
 Leu Leu Leu Lys Glu Asp Val Arg Gly Ser Ala Gln Ser Ser Glu Arg
 10                  15                  20                  25

 AGG GTG GTG GCT CAC ATG CCA GGT GAC ATC ATT ATT GGA GCT CTC TTC
390
 Arg Val Val Ala His Met Pro Gly Asp Ile Ile Ile Gly Ala Leu Phe
                 30                  35                  40

 TCT GTC CAC CAC CAA CCA ACT GTG GAC AAA GTT CAT GAG AGG AAG TGT
438
 Ser Val His His Gln Pro Thr Val Asp Lys Val His Glu Arg Lys Cys
             45                  50                  55

 GGT GCA GTT CGT GAG CAA TAT GGG ATT CAG AGA GTG GAG GCC ATG CTG
486
 Gly Ala Val Arg Glu Gln Tyr Gly Ile Gln Arg Val Glu Ala Met Leu
         60                  65                  70

 CAT ACC TTG GAA AGG ATC AAC TCA GAT CCC ACA CTC TTG CCC AAC ATC
534
 His Thr Leu Glu Arg Ile Asn Ser Asp Pro Thr Leu Leu Pro Asn Ile
     75                  80                  85

 ACA CTT GGC TGT GAG ATC AGA GAT TCC TGC TGG CAT TCC GCT GTG GCC
582
 Thr Leu Gly Cys Glu Ile Arg Asp Ser Cys Trp His Ser Ala Val Ala
 90                  95                  100                 105

 CTA GAA CAA AGC ATT GAG TTT ATA AGG GAT TCC CTC ATC TCT TCA GAA
630
 Leu Glu Gln Ser Ile Glu Phe Ile Arg Asp Ser Leu Ile Ser Ser Glu
                 110                 115                 120

 GAG GAG GAA GGT TTG GTA CGC TGT GTA GAT GGA TCT TCT TCC TTC CGC
678
 Glu Glu Glu Gly Leu Val Arg Cys Val Asp Gly Ser Ser Ser Phe Arg
             125                 130                 135
```

FIGURE 16(A)

```
 TCC AAG AAA CCC ATA GTG GGA GTC ATT GGG CCT GGC TCG AGT TCT GTG
726
 Ser Lys Lys Pro Ile Val Gly Val Ile Gly Pro Gly Ser Ser Ser Val
         140                 145                 150

 GCC ATT CAA GTC CAG AAC TTA CTC CAG CTT TTC AAC ATA CCA CAG ATT
774
 Ala Ile Gln Val Gln Asn Leu Leu Gln Leu Phe Asn Ile Pro Gln Ile
     155                 160                 165

 GCT TAC TCT GCA ACC AGC ATG GAT TTG AGT GAC AAG ACT CTA TTC AAG
822
 Ala Tyr Ser Ala Thr Ser Met Asp Leu Ser Asp Lys Thr Leu Phe Lys
 170                 175                 180                 185

 TAC TTC ATG AGG GTT GTA CCT TCC GAT GCC CAG CAA GCT CGA GCC ATG
870
 Tyr Phe Met Arg Val Val Pro Ser Asp Ala Gln Gln Ala Arg Ala Met
                 190                 195                 200

 GTA GAC ATA GTG AAG AGA TAC AAC TGG ACT TAT GTC TCA GCT GTG CAC
918
 Val Asp Ile Val Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His
             205                 210                 215

 ACA GAA GGC AAC TAT GGA GAA AGT GGG ATG GAG GCT TTC AAA GAT ATG
966
 Thr Glu Gly Asn Tyr Gly Glu Ser Gly Met Glu Ala Phe Lys Asp Met
         220                 225                 230

 TCA GCG AAG GAA GGG ATT TGC ATC GCC CAC TCT TAC AAA ATT TAC AGC
1014
 Ser Ala Lys Glu Gly Ile Cys Ile Ala His Ser Tyr Lys Ile Tyr Ser
     235                 240                 245

 AAT GCT GGG GAA CAG AGC TTT GAC AAG CTG TTG AAA AAA CTC AGA AGT
1062
 Asn Ala Gly Glu Gln Ser Phe Asp Lys Leu Leu Lys Lys Leu Arg Ser
 250                 255                 260                 265

 CAT TTA CCT AAA GCC CGG GTG GTA GCT TGC TTC TGT GAA GGC ATG ACA
1110
 His Leu Pro Lys Ala Arg Val Val Ala Cys Phe Cys Glu Gly Met Thr
                 270                 275                 280

 GTT CGA GGT CTG CTC ATG GCC ATG AGA CGC TTA GGT CTA GCA GGG GAG
1158
 Val Arg Gly Leu Leu Met Ala Met Arg Arg Leu Gly Leu Ala Gly Glu
             285                 290                 295

 TTT CTA CTT CTG GGC AGT GAT GGC TGG GCC GAC AGG TAT GAC GTG ACA
1206
 Phe Leu Leu Leu Gly Ser Asp Gly Trp Ala Asp Arg Tyr Asp Val Thr
         300                 305                 310
```

FIGURE 16(B)

```
 GAT GGA TAT CAG CGA GAA GCT GTC GGT GGA ATT ACA ATC AAG CTT CAG
1254
 Asp Gly Tyr Gln Arg Glu Ala Val Gly Gly Ile Thr Ile Lys Leu Gln
     315                 320                 325

 TCT CCT GAT GTC AAG TGG TTT GAT GAT TAT TAT CTG AAA CTC CGG CCA
1302
 Ser Pro Asp Val Lys Trp Phe Asp Asp Tyr Tyr Leu Lys Leu Arg Pro
 330                 335                 340                 345

 GAA ACA AAC CTC AGA AAC CCT TGG TTT CAA GAA TTT TGG CAG CAT CGT
1350
 Glu Thr Asn Leu Arg Asn Pro Trp Phe Gln Glu Phe Trp Gln His Arg
                 350                 355                 360

 TTT CAG TGC CGG CTG GAA GGG TTT GCA CAG GAG AAC AGC AAG TAC AAC
1398
 Phe Gln Cys Arg Leu Glu Gly Phe Ala Gln Glu Asn Ser Lys Tyr Asn
             365                 370                 375

 AAG ACT TGC AAC AGT TCT CTG ACT CTG AGA ACA CAT CAT GTT CAA GAT
1446
 Lys Thr Cys Asn Ser Ser Leu Thr Leu Arg Thr His His Val Gln Asp
         380                 385                 390

 TCC AAA ATG GGA TTT GTG ATA AAT GCA ATC TAT TCT ATG GCT TAT GGA
1494
 Ser Lys Met Gly Phe Val Ile Asn Ala Ile Tyr Ser Met Ala Tyr Gly
     395                 400                 405

 CTC CAC AAT ATG CAG ATG TCC CTG TGT CCA GGC TAT GCA GGC CTC TGT
1542
 Leu His Asn Met Gln Met Ser Leu Cys Pro Gly Tyr Ala Gly Leu Cys
 410                 415                 420                 425

 GAT GCG ATG AAG CCA ATT GAT GGA CGG AAA CTT TTG GAT TCC CTG ATG
1590
 Asp Ala Met Lys Pro Ile Asp Gly Arg Lys Leu Leu Asp Ser Leu Met
                 430                 435                 440

 AAA ACC AAT TTT ACT GGA GTT TCT GGA GAT ATG ATC CTG TTT GAT GAA
1638
 Lys Thr Asn Phe Thr Gly Val Ser Gly Asp Met Ile Leu Phe Asp Glu
             445                 450                 455

 AAT GGA GAC TCT CCA GGA AGG TAT GAA ATA ATG AAT TTC AAG GAA ATG
1686
 Asn Gly Asp Ser Pro Gly Arg Tyr Glu Ile Met Asn Phe Lys Glu Met
         460                 465                 470

 GGA AAA GAT TAT TTT GAT TAC ATC AAT GTT GGA AGT TGG GAC AAT GGG
1734
 Gly Lys Asp Tyr Phe Asp Tyr Ile Asn Val Gly Ser Trp Asp Asn Gly
     475                 480                 485
```

FIGURE 16(C)

```
 GAA TTA AAA ATG GAT GAT GAC GAA GTG TGG TCC AAG AAA AAT AAC ATC
1782
 Glu Leu Lys Met Asp Asp Asp Glu Val Trp Ser Lys Lys Asn Asn Ile
 490                 495                 500                 505

 ATC AGA TCT GTG TGC AGT GAA CCA TGT GAG AAA GGC CAA ATA AAG GTG
1830
 Ile Arg Ser Val Cys Ser Glu Pro Cys Glu Lys Gly Gln Ile Lys Val
                 510                 515                 520

 ATC CGG AAG GGA GAA GTC AGC TGT TGT TGG ACC TGC ACA CCT TGT AAA
1878
 Ile Arg Lys Gly Glu Val Ser Cys Cys Trp Thr Cys Thr Pro Cys Lys
             525                 530                 535

 GAG AAT GAG TAT GTG TTT GAT GAG TAC ACC TGC AAG GCA TGC CAA CTG
1926
 Glu Asn Glu Tyr Val Phe Asp Glu Tyr Thr Cys Lys Ala Cys Gln Leu
         540                 545                 550

 GGC TCC TGG CCC ACT GAC GAC CTG ACA GGC TGT GAC TTG ATC CCA GTC
1974
 Gly Ser Trp Pro Thr Asp Asp Leu Thr Gly Cys Asp Leu Ile Pro Val
     555                 560                 565

 CAG TAT CTT CGC TGG GGT GAC CCT GAG CCG ATT GCA GCT GTG GTG TTT
2022
 Gln Tyr Leu Arg Trp Gly Asp Pro Glu Pro Ile Ala Ala Val Val Phe
 570                 575                 580                 585

 GCC TGC CTC GGT CTG CTA GCC ACC TTA TTC GTT ACT GTA ATC TTC ATC
2070
 Ala Cys Leu Gly Leu Leu Ala Thr Leu Phe Val Thr Val Ile Phe Ile
                 590                 595                 600

 ATT TAT CGG GAC ACT CCG GTG GTC AAG TCC TCC AGC AGG GAG CTC TGC
2118
 Ile Tyr Arg Asp Thr Pro Val Val Lys Ser Ser Ser Arg Glu Leu Cys
             605                 610                 615

 TAC ATT ATC CTT GCT GGC ATC TGC CTG GGA TAC TTA TGT ACC TTT TGC
2166
 Tyr Ile Ile Leu Ala Gly Ile Cys Leu Gly Tyr Leu Cys Thr Phe Cys
         620                 625                 630

 CTC ATT GCA AAG CCC AAG CAG ATT TAC TGC TAT CTT CAG AGA ATT GGC
2214
 Leu Ile Ala Lys Pro Lys Gln Ile Tyr Cys Tyr Leu Gln Arg Ile Gly
     635                 640                 645

 ATT GGA CTC TCT CCA GCC ATG AGC TAC TCA GCC CTT GTA ACC AAG ACC
2262
 Ile Gly Leu Ser Pro Ala Met Ser Tyr Ser Ala Leu Val Thr Lys Thr
 650                 655                 660                 665
```

FIGURE 16(D)

```
 AAC CGT ATT GCA AGG ATT CTA GCT GGA AGC AAG AAG AAA ATC TGT ACC
2310
 Asn Arg Ile Ala Arg Ile Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr
                 670                 675                 680

 AAG AAG CCC AGA TTC ATG AGC GCC TGT GCT CAG TTA GTG ATC GCT TTC
2358
 Lys Lys Pro Arg Phe Met Ser Ala Cys Ala Gln Leu Val Ile Ala Phe
             685                 690                 695

 ATT CTC ATC TGT ATT CAG CTG GGC ATT ATC GTG GCC CTC TTT ATC ATG
2406
 Ile Leu Ile Cys Ile Gln Leu Gly Ile Ile Val Ala Leu Phe Ile Met
         700                 705                 710

 GAG CCT CCG GAT ATA ATG CAT GAC TAT CCA AGC ATC CGA GAA GTC TAC
2454
 Glu Pro Pro Asp Ile Met His Asp Tyr Pro Ser Ile Arg Glu Val Tyr
     715                 720                 725

 TTG ATT TGT AAC ACC ACC AAC CTA GGG GTT GTT ACT CCT CTT GGA TAC
2502
 Leu Ile Cys Asn Thr Thr Asn Leu Gly Val Val Thr Pro Leu Gly Tyr
 730                 735                 740                 745

 AAT GGA TTA TTG ATT TTG AGT TGC ACG TTC TAT GCG TTT AAG ACC AGA
2550
 Asn Gly Leu Leu Ile Leu Ser Cys Thr Phe Tyr Ala Phe Lys Thr Arg
                 750                 755                 760

 AAT GTT CCA GCC AAC TTT AAT GAG GCC AAA TAT ATT GCT TTC ACC ATG
2598
 Asn Val Pro Ala Asn Phe Asn Glu Ala Lys Tyr Ile Ala Phe Thr Met
             765                 770                 775

 TAC ACA ACC TGC ATC ATA TGG CTG GCC TTC GTG CCT ATC TAC TTT GGC
2646
 Tyr Thr Thr Cys Ile Ile Trp Leu Ala Phe Val Pro Ile Tyr Phe Gly
         780                 785                 790

 AGC AAC TAC AAA ATC ATC ACC ATG TGT TTT TCA GTC AGC CTC AGT GCC
2694
 Ser Asn Tyr Lys Ile Ile Thr Met Cys Phe Ser Val Ser Leu Ser Ala
     795                 800                 805

 ACA GTG GCC CTG GGT TGC ATG TTT GTC CCG AAG GTG TAC ATC ATC CTA
2742
 Thr Val Ala Leu Gly Cys Met Phe Val Pro Lys Val Tyr Ile Ile Leu
 810                 815                 820                 825

 GCC AAA CCG GAG AGA AAT GTG CGC AGC GCC TTC ACA ACC TCT ACA GTG
2790
 Ala Lys Pro Glu Arg Asn Val Arg Ser Ala Phe Thr Thr Ser Thr Val
                 830                 835                 840
```

FIGURE 16(E)

```
 GTA CGC ATG CAT GTA GGA GAC GGC AAA TCA TCG TCC GCT GCC AGC AGA
2838
 Val Arg Met His Val Gly Asp Gly Lys Ser Ser Ser Ala Ala Ser Arg
            845                 850                 855

 TCC AGC AGC CTA GTC AAC CTG TGG AAG AGG AGG GGC TCC TCT GGG GAA
2886
 Ser Ser Ser Leu Val Asn Leu Trp Lys Arg Arg Gly Ser Ser Gly Glu
        860                 865                 870

 ACC CTA AGC TCC AAC GGA AAA TCT GTG ACT TGG GCC CAG AAT GAG AAG
2934
 Thr Leu Ser Ser Asn Gly Lys Ser Val Thr Trp Ala Gln Asn Glu Lys
    875                 880                 885

 AGT ACC CGG GGG CAA CAT TTG TGG CAG CGA CTG TCT GTC CAC ATC AAC
2982
 Ser Thr Arg Gly Gln His Leu Trp Gln Arg Leu Ser Val His Ile Asn
 890                 895                 900                 905

 AAG AAG GAG AAC CCC AAC CAA ACG GCG GTC ATC AAA CCC TTT CCC AAG
3030
 Lys Lys Glu Asn Pro Asn Gln Thr Ala Val Ile Lys Pro Phe Pro Lys
                910                 915                 920

 AGC ACA GAG AAC CGG GGG CCT GGT GCA GCG GCA GGT GGT GGC TCG GGT
3078
 Ser Thr Glu Asn Arg Gly Pro Gly Ala Ala Ala Gly Gly Gly Ser Gly
            925                 930                 935

 CCC GGT GTA GCT GGT GCT GGT AAC GCA GGA TGC ACA GCA ACA GGT GGC
3126
 Pro Gly Val Ala Gly Ala Gly Asn Ala Gly Cys Thr Ala Thr Gly Gly
        940                 945                 950

 CCG GAG CCA CCG GAC GCC GGT CCC AAA GCG CTT TAT GAT GTG GCG GAG
3174
 Pro Glu Pro Pro Asp Ala Gly Pro Lys Ala Leu Tyr Asp Val Ala Glu
    955                 960                 965

 GCA GAG GAG AGC TTC CCC GCG GCT GCC AGG CCG CGC TCG CCA TCG CCC
3222
 Ala Glu Glu Ser Phe Pro Ala Ala Ala Arg Pro Arg Ser Pro Ser Pro
 970                 975                 980                 985

 ATC AGC ACG TTG AGC CAC CTC GCA GGC TCG GCG GGC CGC ACA GAT GAT
3270
 Ile Ser Thr Leu Ser His Leu Ala Gly Ser Ala Gly Arg Thr Asp Asp
                990                 995                1000

 GAT GCG CCG TCG CTG CAC TCT GAG ACA GCT GCT CGC AGC AGC TCG TCC
3318
 Asp Ala Pro Ser Leu His Ser Glu Thr Ala Ala Arg Ser Ser Ser Ser
            1005                1010                1015
```

FIGURE 16(F)

```
 CAG GGC TCG CTC ATG GAG CAG ATC AGC AGC GTA GTG ACG CGC TTC ACA
3366
 Gln Gly Ser Leu Met Glu Gln Ile Ser Ser Val Val Thr Arg Phe Thr
        1020                1025                1030

 GCC AAC ATC TCC GAG CTC AAC TCC ATG ATG TTG TCC ACC GCG GCC ACG
3414
 Ala Asn Ile Ser Glu Leu Asn Ser Met Met Leu Ser Thr Ala Ala Thr
    1035                1040                1045

 CCG GGG CCC CCT GGT ACT CCA ATC TGC TCC TCC TAC CTG ATC CCC AAA
3462
 Pro Gly Pro Pro Gly Thr Pro Ile Cys Ser Ser Tyr Leu Ile Pro Lys
1050                1055                1060                1065

 GAG ATT CAG CTG CCC ACG ACC ATG ACG ACC TTC GCA GAG ATC CAG CCG
3510
 Glu Ile Gln Leu Pro Thr Thr Met Thr Thr Phe Ala Glu Ile Gln Pro
                1070                1075                1080

 CTG CCG GCC ATC GAG GTG ACC GGA GGA GCT CAG GGG GCG ACA GGC GTA
3558
 Leu Pro Ala Ile Glu Val Thr Gly Gly Ala Gln Gly Ala Thr Gly Val
            1085                1090                1095

 TCA CCT GCC CAG GAG ACG CCC ACA GGA GCT GAA TCC GCC CCG GGC AAA
3606
 Ser Pro Ala Gln Glu Thr Pro Thr Gly Ala Glu Ser Ala Pro Gly Lys
        1100                1105                1110

 CCA GAT CTG GAG GAG CTT GTG GCC CTC ACT CCA CCA TCG CCC TTC AGG
3654
 Pro Asp Leu Glu Glu Leu Val Ala Leu Thr Pro Pro Ser Pro Phe Arg
    1115                1120                1125

 GAC TCG GTG GAC TCG GGG AGC ACC ACC CCA AAC TCT CCA GTC TCA GAA
3702
 Asp Ser Val Asp Ser Gly Ser Thr Thr Pro Asn Ser Pro Val Ser Glu
1130                1135                1140                1145

 TCG GCC CTC TGC ATC CCA TCC TCT CCC AAA TAT GAC ACT CTC ATC ATC
3750
 Ser Ala Leu Cys Ile Pro Ser Ser Pro Lys Tyr Asp Thr Leu Ile Ile
                1150                1155                1160

 AGA GAT TAC ACG CAG AGT TCT TCA TCG TTG TGAGCCACTG GAAACTTCCC TGA
3803
 Arg Asp Tyr Thr Gln Ser Ser Ser Ser Leu
            1165                1170

 GATAGGCCTG TGCAGCAGAC CTCGGTGTTC CAGATGGGCG CAGCGACACC CATAGTCACC
3863
 CCAACACAGA ACCCGAGGTG AAAAATGCCA AGGACACACA CGGTGGAGAC ATGAG
3919
```

FIGURE 16(G)

human mGluR5

```
Met Val Leu Leu Leu Ile Leu Ser Val Leu Leu Leu Lys Glu Asp Val
 1               5                  10                  15
Arg Gly Ser Ala Gln Ser Ser Glu Arg Arg Val Val Ala His Met Pro
            20                  25                  30
Gly Asp Ile Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro Thr
        35                  40                  45
Val Asp Lys Val His Glu Arg Lys Cys Gly Ala Val Arg Glu Gln Tyr
    50                  55                  60
Gly Ile Gln Arg Val Glu Ala Met Leu His Thr Leu Glu Arg Ile Asn
65                  70                  75                  80
Ser Asp Pro Thr Leu Leu Pro Asn Ile Thr Leu Gly Cys Glu Ile Arg
                85                  90                  95
Asp Ser Cys Trp His Ser Ala Val Ala Leu Glu Gln Ser Ile Glu Phe
            100                 105                 110
Ile Arg Asp Ser Leu Ile Ser Ser Glu Glu Glu Glu Gly Leu Val Arg
        115                 120                 125
Cys Val Asp Gly Ser Ser Ser Ser Phe Arg Ser Lys Lys Pro Ile Val
    130                 135                 140
Gly Val Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val Gln Asn
145                 150                 155                 160
Leu Leu Gln Leu Phe Asn Ile Pro Gln Ile Ala Tyr Ser Ala Thr Ser
                165                 170                 175
Met Asp Leu Ser Asp Lys Thr Leu Phe Lys Tyr Phe Met Arg Val Val
            180                 185                 190
Pro Ser Asp Ala Gln Gln Ala Arg Ala Met Val Asp Ile Val Lys Arg
        195                 200                 205
Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn Tyr Gly
    210                 215                 220
Glu Ser Gly Met Glu Ala Phe Lys Asp Met Ser Ala Lys Glu Gly Ile
225                 230                 235                 240
Cys Ile Ala His Ser Tyr Lys Ile Tyr Ser Asn Ala Gly Glu Gln Ser
                245                 250                 255
Phe Asp Lys Leu Leu Lys Lys Leu Thr Ser His Leu Pro Lys Ala Arg
            260                 265                 270
Val Val Ala Cys Phe Cys Glu Gly Met Thr Val Arg Gly Leu Leu Met
        275                 280                 285
Ala Met Arg Arg Leu Gly Leu Ala Gly Glu Phe Leu Leu Leu Gly Ser
    290                 295                 300
Asp Gly Trp Ala Asp Arg Tyr Asp Val Thr Asp Gly Tyr Gln Arg Glu
305                 310                 315                 320
Ala Val Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro Asp Val Lys Trp
                325                 330                 335
Phe Asp Asp Tyr Tyr Leu Lys Leu Arg Pro Glu Thr Asn His Arg Asn
            340                 345                 350
Pro Trp Phe Gln Glu Phe Trp Gln His Arg Phe Gln Cys Arg Leu Glu
        355                 360                 365
Gly Phe Pro Gln Glu Asn Ser Lys Tyr Asn Lys Thr Cys Asn Ser Ser
    370                 375                 380
Leu Thr Leu Lys Thr His His Val Gln Asp Ser Lys Met Gly Phe Val
385                 390                 395                 400
Ile Asn Ala Ile Tyr Ser Met Ala Tyr Gly Leu His Asn Met Gln Met
                405                 410                 415
Ser Leu Cys Pro Gly Tyr Ala Gly Leu Cys Asp Ala Met Lys Pro Ile
            420                 425                 430
```

FIGURE 17(A)

```
Asp Gly Arg Lys Leu Leu Glu Ser Leu Met Lys Thr Asn Phe Thr Gly
          435                 440                 445
Val Ser Gly Asp Thr Ile Leu Phe Asp Glu Asn Gly Asp Ser Pro Gly
      450                 455                 460
Arg Tyr Glu Ile Met Asn Phe Lys Glu Met Gly Lys Asp Tyr Phe Asp
465                 470                 475                 480
Tyr Ile Asn Val Gly Ser Trp Asp Asn Gly Glu Leu Lys Met Asp Asp
                485                 490                 495
Asp Glu Val Trp Ser Lys Lys Ser Asn Ile Ile Arg Ser Val Cys Ser
            500                 505                 510
Glu Pro Cys Glu Lys Gly Gln Ile Lys Val Ile Arg Lys Gly Glu Val
        515                 520                 525
Ser Cys Cys Trp Thr Cys Thr Pro Cys Lys Glu Asn Glu Tyr Val Phe
    530                 535                 540
Asp Glu Tyr Thr Cys Lys Ala Cys Gln Leu Gly Ser Trp Pro Thr Asp
545                 550                 555                 560
Asp Leu Thr Gly Cys Asp Leu Ile Pro Val Gln Tyr Leu Arg Trp Gly
                565                 570                 575
Asp Pro Glu Pro Ile Ala Ala Val Val Phe Ala Cys Leu Gly Leu Leu
            580                 585                 590
Ala Thr Leu Phe Val Thr Val Val Phe Ile Ile Tyr Arg Asp Thr Pro
        595                 600                 605
Val Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile Ile Leu Ala Gly
    610                 615                 620
Ile Cys Leu Gly Tyr Leu Cys Thr Phe Cys Leu Ile Ala Lys Pro Lys
625                 630                 635                 640
Gln Ile Tyr Cys Tyr Leu Gln Arg Ile Gly Ile Gly Leu Ser Pro Ala
                645                 650                 655
Met Ser Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala Arg Ile
            660                 665                 670
Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr Lys Lys Pro Arg Phe Met
        675                 680                 685
Ser Ala Cys Ala Gln Leu Val Ile Ala Phe Ile Leu Ile Cys Ile Gln
    690                 695                 700
Leu Gly Ile Ile Val Ala Leu Phe Ile Met Glu Pro Pro Asp Ile Met
705                 710                 715                 720
His Asp Tyr Pro Ser Ile Arg Glu Val Tyr Leu Ile Cys Asn Thr Thr
                725                 730                 735
Asn Leu Gly Val Val Thr Pro Leu Gly Tyr Asn Gly Leu Leu Ile Leu
            740                 745                 750
Ser Cys Thr Phe Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala Asn Phe
        755                 760                 765
Asn Glu Ala Lys Tyr Ile Ala Phe Thr Met Tyr Thr Thr Cys Ile Ile
    770                 775                 780
Trp Leu Ala Phe Val Pro Ile Tyr Phe Gly Ser Asn Tyr Lys Ile Ile
785                 790                 795                 800
Thr Met Cys Phe Ser Val Ser Leu Ser Ala Thr Val Ala Leu Gly Cys
                805                 810                 815
Met Phe Val Pro Lys Val Tyr Ile Ile Leu Ala Lys Pro Glu Arg Asn
            820                 825                 830
Val Arg Ser Ala Phe Thr Thr Ser Thr Val Val Arg Met His Val Gly
        835                 840                 845
Asp Gly Lys Ser Ser Ser Ala Ala Ser Arg Ser Ser Ser Leu Val Asn
    850                 855                 860
Leu Trp Lys Arg Arg Gly Ser Ser Gly Glu Thr Leu Arg Tyr Lys Asp
```

FIGURE 17(B)

```
865                 870                 875                 880
Arg Arg Leu Ala Gln His Lys Ser Glu Ile Glu Cys Phe Thr Pro Lys
                885                 890                 895
Gly Ser Met Gly Asn Gly Gly Arg Ala Thr Met Ser Ser Ser Asn Gly
            900                 905                 910
Lys Ser Val Thr Trp Ala Gln Asn Glu Lys Ser Ser Arg Gly Gln His
        915                 920                 925
Leu Trp Gln Arg Leu Ser Ile His Ile Asn Lys Lys Glu Asn Pro Asn
    930                 935                 940
Gln Thr Ala Val Ile Lys Pro Phe Pro Lys Ser Thr Glu Ser Arg Gly
945                 950                 955                 960
Leu Gly Ala Gly Ala Gly Ala Gly Gly Ser Ala Gly Gly Val Gly Ala
                965                 970                 975
Thr Gly Gly Ala Gly Cys Ala Gly Ala Gly Pro Gly Gly Pro Glu Ser
            980                 985                 990
Pro Asp Ala Gly Pro Lys Ala Leu Tyr Asp Val Ala Glu Ala Glu Glu
        995                1000                1005
His Phe Pro Ala Pro Ala Arg Pro Arg Ser Pro Ser Pro Ile Ser Thr
   1010                1015                1020
Leu Ser His Arg Ala Gly Ser Ala Ser Arg Thr Asp Asp Asp Val Pro
025                1030                1035                1040
Ser Leu His Ser Glu Pro Val Ala Arg Ser Ser Ser Ser Gln Gly Ser
               1045                1050                1055
Leu Met Glu Gln Ile Ser Ser Val Val Thr Arg Phe Thr Ala Asn Ile
           1060                1065                1070
Ser Glu Leu Asn Ser Met Met Leu Ser Thr Ala Ala Pro Ser Pro Gly
       1075                1080                1085
Val Gly Ala Pro Leu Cys Ser Ser Tyr Leu Ile Pro Lys Glu Ile Gln
   1090                1095                1100
Leu Pro Thr Thr Met Thr Thr Phe Ala Glu Ile Gln Pro Leu Pro Ala
105                1110                1115                1120
Ile Glu Val Thr Gly Gly Ala Gln Pro Ala Ala Gly Ala Gln Ala Ala
               1125                1130                1135
Gly Asp Ala Ala Arg Glu Ser Pro Ala Ala Gly Pro Glu Ala Ala Ala
           1140                1145                1150
Ala Lys Pro Asp Leu Glu Glu Leu Val Ala Leu Thr Pro Pro Ser Pro
       1155                1160                1165
Phe Arg Asp Ser Val Asp Ser Gly Ser Thr Thr Pro Asn Ser Pro Val
   1170                1175                1180
Ser Glu Ser Ala Leu Cys Ile Pro Ser Ser Pro Lys Tyr Asp Thr Leu
185                1190                1195                1200
Ile Ile Arg Asp Tyr Thr Gln Ser Ser Ser Ser Leu
               1205                1210
```

FIGURE 17(C)

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Johnson, Thomas E.
Sikela, James M.
Simpson, Victoria J.
Rikke, Brad A.

(ii) TITLE OF THE INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES ASSOCIATED WITH CNS DEPRESSANT SENSITIVITY AND METHODS OF USE THEREOF

(iii) NUMBER OF SEQUENCES: 7

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: MORRISON & FOERSTER
(B) STREET: 755 PAGE MILL ROAD
(C) CITY: Palo Alto
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94304-1018

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: To Be Assigned
(B) FILING DATE: Herewith
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 60/073,190
(B) FILING DATE: January 30, 1998

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Catherine, Polizzi M
(B) REGISTRATION NUMBER: 40,130
(C) REFERENCE/DOCKET NUMBER: 34293-20003.00

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 650-813-5600
(B) TELEFAX: 650-494-0792
(C) TELEX: 706141

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3591 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 1...3591
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
ATG GTC CTT CTG TTG ATT CTG TCA GTC CTA CTT CTG AAA GAA GAT GTA      48
Met Val Leu Leu Leu Ile Leu Ser Val Leu Leu Leu Lys Glu Asp Val
 1               5                  10                  15

CGA GGA AGT GCA CAG TCC AGT GAG AGG AGG GTG GTG GCT CAC ATG CCA      96
Arg Gly Ser Ala Gln Ser Ser Glu Arg Arg Val Val Ala His Met Pro
            20                  25                  30

GGT GAC ATT ATT ATT GGA GCT CTC TTC TCG GTC CAC CAC CAA CCA ACT     144
Gly Asp Ile Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro Thr
        35                  40                  45

GTG GAC AAA GTT CAT GAG AGG AAG TGT GGG GCA GTC CGT GAG CAG TAT     192
Val Asp Lys Val His Glu Arg Lys Cys Gly Ala Val Arg Glu Gln Tyr
    50                  55                  60

GGC ATT CAG AGA GTG GAA GCC ATG CTG CAT ACC TTG GAA AGG ATC AAT     240
Gly Ile Gln Arg Val Glu Ala Met Leu His Thr Leu Glu Arg Ile Asn
65                  70                  75                  80

TCA GAT CCC ACA CTC TTG CCC AAC ATC ACA CTA GGC TGT GAG ATA AGA     288
Ser Asp Pro Thr Leu Leu Pro Asn Ile Thr Leu Gly Cys Glu Ile Arg
                85                  90                  95

GAT TCC TGC TGG CAT TCT GCT GTG GCC CTA GAG CAA AGC ATT GAG TTT     336
Asp Ser Cys Trp His Ser Ala Val Ala Leu Glu Gln Ser Ile Glu Phe
            100                 105                 110

ATA AGG GAT TCC CTC ATC TCT TCG GAA GAG GAA GAA GGC TTG GTA CGC     384
Ile Arg Asp Ser Leu Ile Ser Ser Glu Glu Glu Glu Gly Leu Val Arg
        115                 120                 125

TGT GTA GAT GGC TCT TCT TCC TTC CGC TCC AAG AAA CCC ATA GTG GGA     432
Cys Val Asp Gly Ser Ser Ser Phe Arg Ser Lys Lys Pro Ile Val Gly
    130                 135                 140

GTC ATT GGG CCT GGC TCG AGT TCT GTG GCC ATT CAA GTT CAG AAC TTG     480
Val Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val Gln Asn Leu
145                 150                 155                 160

CTC CAG CTT TTC AAC ATA CCT CAG ATT GCT TAC TCT GCA ACT AGC ATG     528
Leu Gln Leu Phe Asn Ile Pro Gln Ile Ala Tyr Ser Ala Thr Ser Met
                165                 170                 175

GAT TTG AGT GAC AAG ACT CTA TTC AAG TAC TTC ATG AGG GTT GTA CCT     576
Asp Leu Ser Asp Lys Thr Leu Phe Lys Tyr Phe Met Arg Val Val Pro
            180                 185                 190

TCG GAT GCC CAG CAA GCC CGA GCC ATG GTA GAC ATA GTG AAG AGA TAC     624
Ser Asp Ala Gln Gln Ala Arg Ala Met Val Asp Ile Val Lys Arg Tyr
        195                 200                 205

AAC TGG ACT TAT GTC TCA GCT GTG CAC ACA GAA GGC AAC TAT GGA GAA     672
Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn Tyr Gly Glu
    210                 215                 220

AGT GGG ATG GAG GCT TTC AAA GAT ATG TCA GCG AAG GAA GGG ATT TGC     720
Ser Gly Met Glu Ala Phe Lys Asp Met Ser Ala Lys Glu Gly Ile Cys
```

225 230 235 240

ATC GCC CAC TCT TAC AAA ATC TAC AGC AAT GCT GGG GAA CAG AGC TTT 768
Ile Ala His Ser Tyr Lys Ile Tyr Ser Asn Ala Gly Glu Gln Ser Phe
245 250 255

GAC AAG CTG TTG AAA AAG CTC AGA AGT CAT TTA CCT AAA GCC CGG GTG 816
Asp Lys Leu Leu Lys Lys Leu Arg Ser His Leu Pro Lys Ala Arg Val
260 265 270

GTA GCC TGC TTC TGT GAA GGC ATG ACA GTT CGA GGT CTG CTC ATG GCC 864
Val Ala Cys Phe Cys Glu Gly Met Thr Val Arg Gly Leu Leu Met Ala
275 280 285

ATG AGA CGC TTG GGT CTA GCA GGG GAA TTT CTA CTT CTG GGC AGT GAT 912
Met Arg Arg Leu Gly Leu Ala Gly Glu Phe Leu Leu Leu Gly Ser Asp
290 295 300

GGC TGG GCT GAC AGG TAT GAC GTG ACA GAT GGG TAT CAG CGA GAA GCT 960
Gly Trp Ala Asp Arg Tyr Asp Val Thr Asp Gly Tyr Gln Arg Glu Ala
305 310 315 320

GTC GGT GGG ATT ACA ATC AAG CTC CAG TCT CCT GAT GTC AAG TGG TTT 1008
Val Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro Asp Val Lys Trp Phe
325 330 335

GAT GAT TAT TAT CTG AAG CTC CGG CCA GAA ACA AAC CTC AGA AAC CCT 1056
Asp Asp Tyr Tyr Leu Lys Leu Arg Pro Glu Thr Asn Leu Arg Asn Pro
340 345 350

TGG TTT CAA GAA TTT TGG CAG CAT CGT TTT CAG TGC CGG CTA GAA GGG 1104
Trp Phe Gln Glu Phe Trp Gln His Arg Phe Gln Cys Arg Leu Glu Gly
355 360 365

TTT GCA CAG GAG AAC AGC AAG TAC AAC AAG ACT TGC AAC AGT TCT CTA 1152
Phe Ala Gln Glu Asn Ser Lys Tyr Asn Lys Thr Cys Asn Ser Ser Leu
370 375 380

ACT CTG AGA ACG CAT CAT GTT CAA GAT TCC AAA ATG GGA TTT GTG ATC 1200
Thr Leu Arg Thr His His Val Gln Asp Ser Lys Met Gly Phe Val Ile
385 390 395 400

AAT GCA ATC TAT TCT ATG GCT TAT GGG CTC CAC AAC ATG CAG ATG TCC 1248
Asn Ala Ile Tyr Ser Met Ala Tyr Gly Leu His Asn Met Gln Met Ser
405 410 415

CTG TGT CCA GGC TAT GCA GGC CTC TGT GAT GCA ATG AAG CCA ATT GAT 1296
Leu Cys Pro Gly Tyr Ala Gly Leu Cys Asp Ala Met Lys Pro Ile Asp
420 425 430

GGG CGG AAA CTT TTG GAC TCC CTG ATG AAA ACC AAC TTT ACT GGA GTT 1344
Gly Arg Lys Leu Leu Asp Ser Leu Met Lys Thr Asn Phe Thr Gly Val
435 440 445

TCC GGA GAT ATG ATT CTA TTT GAT GAA AAT GGA GAC TCT CCA GGA AGG 1392
Ser Gly Asp Met Ile Leu Phe Asp Glu Asn Gly Asp Ser Pro Gly Arg
450 455 460

TAT GAA ATA ATG AAT TTC AAG GAA ATG GGA AAA GAT TAT TTT GAT TAC 1440
Tyr Glu Ile Met Asn Phe Lys Glu Met Gly Lys Asp Tyr Phe Asp Tyr
465 470 475 480

```
ATC AAT GTT GGA AGT TGG GAC AAT GGG GAA TTA AAA ATG GAT GAT GAT     1488
Ile Asn Val Gly Ser Trp Asp Asn Gly Glu Leu Lys Met Asp Asp Asp
                485                 490                 495

GAA GTG TGG TCC AAG AAA AAT AAC ATC ATC AGA TCT GTG TGC AGT GAA     1536
Glu Val Trp Ser Lys Lys Asn Asn Ile Ile Arg Ser Val Cys Ser Glu
            500                 505                 510

CCG TGT GAG AAA GGA CAG ATA AAG GTG ATC CGG AAG GGA GAA GTC AGC     1584
Pro Cys Glu Lys Gly Gln Ile Lys Val Ile Arg Lys Gly Glu Val Ser
        515                 520                 525

TGT TGT TGG ACC TGC ACA CCT TGT AAG GAG AAT GAG TAT GTT TTT GAT     1632
Cys Cys Trp Thr Cys Thr Pro Cys Lys Glu Asn Glu Tyr Val Phe Asp
    530                 535                 540

GAG TAC ACC TGC AAG GCG TGC CAA CTG GGG TCC TGG CCC ACT GAC GAC     1680
Glu Tyr Thr Cys Lys Ala Cys Gln Leu Gly Ser Trp Pro Thr Asp Asp
545                 550                 555                 560

TTG ACA GGT TGT GAT TTG ATC CCA GTC CAG TAT CTT CGA TGG GGT GAC     1728
Leu Thr Gly Cys Asp Leu Ile Pro Val Gln Tyr Leu Arg Trp Gly Asp
                565                 570                 575

CCT GAG CCC ATT GCA GCT GTG GTG TTT GCC TGC CTC GGT CTG CTA GCC     1776
Pro Glu Pro Ile Ala Ala Val Val Phe Ala Cys Leu Gly Leu Leu Ala
            580                 585                 590

ACC CTC TTC GTT ACT GTA ATC TTC ATC ATT TAT CGG GAC ACT CCA GTG     1824
Thr Leu Phe Val Thr Val Ile Phe Ile Ile Tyr Arg Asp Thr Pro Val
        595                 600                 605

GTC AAG TCT TCC AGC AGG GAA CTC TGC TAC ATT ATC CTT GCT GGC ATC     1872
Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile Ile Leu Ala Gly Ile
    610                 615                 620

TGC CTG GGT TAC TTG TGT ACC TTC TGC CTC ATT GCA AAG CCC AAA CAG     1920
Cys Leu Gly Tyr Leu Cys Thr Phe Cys Leu Ile Ala Lys Pro Lys Gln
625                 630                 635                 640

ATT TAC TGC TAC CTT CAG AGA ATT GGC ATC GGT CTC TCT CCA GCC ATG     1968
Ile Tyr Cys Tyr Leu Gln Arg Ile Gly Ile Gly Leu Ser Pro Ala Met
                645                 650                 655

AGC TAC TCA GCC CTT GTA ACG AAG ACC AAC CGT ATT GCA AGG ATT CTA     2016
Ser Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala Arg Ile Leu
            660                 665                 670

GCT GGC AGC AAG AAG AAG ATC TGT ACC AAA AAG CCC AGA TTC ATG AGC     2064
Ala Gly Ser Lys Lys Lys Ile Cys Thr Lys Lys Pro Arg Phe Met Ser
        675                 680                 685

GCC TGT GCT CAG TTA GTG ATC GCT TTC ATT CTC ATC TGT ATA CAG TTG     2112
Ala Cys Ala Gln Leu Val Ile Ala Phe Ile Leu Ile Cys Ile Gln Leu
    690                 695                 700

GGT ATT ATT GTG GCC CTC TTT ATC ATG GAG CCT CCG GAT ATA ATG CAT     2160
Gly Ile Ile Val Ala Leu Phe Ile Met Glu Pro Pro Asp Ile Met His
705                 710                 715                 720
```

```
GAC TAT CCA AGC ATC CGA GAA GTC TAC TTG ATT TGT AAC ACC ACC AAC   2208
Asp Tyr Pro Ser Ile Arg Glu Val Tyr Leu Ile Cys Asn Thr Thr Asn
                725                 730                 735

CTA GGG GTT GTC ACT CCT CTT GGA TAC AAT GGA TTA TTG ATT TTG AGT   2256
Leu Gly Val Val Thr Pro Leu Gly Tyr Asn Gly Leu Leu Ile Leu Ser
            740                 745                 750

TGC ACA TTC TAT GCG TTC AAG ACC AGA AAT GTT CCA GCC AAC TTT AAC   2304
Cys Thr Phe Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala Asn Phe Asn
        755                 760                 765

GAG GCC AAA TAT ATT GCT TTC ACC ATG TAC ACA ACC TGC ATC ATA TGG   2352
Glu Ala Lys Tyr Ile Ala Phe Thr Met Tyr Thr Thr Cys Ile Ile Trp
    770                 775                 780

CTG GCC TTT GTG CCT ATC TAC TTT GGC AGC AAC TAC AAA ATC ATC ACC   2400
Leu Ala Phe Val Pro Ile Tyr Phe Gly Ser Asn Tyr Lys Ile Ile Thr
785                 790                 795                 800

ATG TGT TTC TCA GTC AGC CTC AGT GCC ACA GTG GCC CTG GGT TGC ATG   2448
Met Cys Phe Ser Val Ser Leu Ser Ala Thr Val Ala Leu Gly Cys Met
                805                 810                 815

TTT GTG CCG AAG GTG TAC ATC ATC CTA GCC AAA CCG GAG AGA AAT GTG   2496
Phe Val Pro Lys Val Tyr Ile Ile Leu Ala Lys Pro Glu Arg Asn Val
            820                 825                 830

CGC AGC GCC TTC ACA ACC TCT ACA GTG GTG CGC ATG CAC GTA GGA GAT   2544
Arg Ser Ala Phe Thr Thr Ser Thr Val Val Arg Met His Val Gly Asp
        835                 840                 845

GGC AAG TCA TCA TCC GCT GCC AGC AGA TCC AGC AGC CTA GTC AAC CTG   2592
Gly Lys Ser Ser Ser Ala Ala Ser Arg Ser Ser Ser Leu Val Asn Leu
    850                 855                 860

TGG AAG AGG AGG GGC TCG TCT GGG GAA ACC CTA AGG TAC AAA GAC AGG   2640
Trp Lys Arg Arg Gly Ser Ser Gly Glu Thr Leu Arg Tyr Lys Asp Arg
865                 870                 875                 880

AGA CTG GCC CAG CAC AAG TCG GAA ATA GAG TGT TTC ACC CCC AAA GGG   2688
Arg Leu Ala Gln His Lys Ser Glu Ile Glu Cys Phe Thr Pro Lys Gly
                885                 890                 895

AGT ATG GGG AAT GGT GGG AGA GCA ACA ATG AGC AGC TCC AAT GGA AAA   2736
Ser Met Gly Asn Gly Gly Arg Ala Thr Met Ser Ser Ser Asn Gly Lys
            900                 905                 910

TCC GTG ACT TGG GCC CAG AAT GAG AAG AGC ACC CGG GGG CAG CAC CTG   2784
Ser Val Thr Trp Ala Gln Asn Glu Lys Ser Thr Arg Gly Gln His Leu
        915                 920                 925

TGG CAG CGA CTG TCT GTC CAT ATC AAC AAG AAG GAG AAC CCC AAC CAG   2832
Trp Gln Arg Leu Ser Val His Ile Asn Lys Lys Glu Asn Pro Asn Gln
    930                 935                 940

ACA GCA GTC ATC AAA CCC TTC CCC AAG AGC ACA GAG AGC CGC GGG CAG   2880
Thr Ala Val Ile Lys Pro Phe Pro Lys Ser Thr Glu Ser Arg Gly Gln
945                 950                 955                 960

GGT GCA GGG GCA GGT GGT GGC TCT GGC CCC GGT GCA GCT GGT GCT GGT   2928
```

```
Gly Ala Gly Ala Gly Gly Gly Ser Gly Pro Gly Ala Ala Gly Ala Gly
                965                 970                 975

AGC GCA GGA TGC ACA GCG ACA GGC GGC CCA GAG CCA CCA GAC GCC GGC    2976
Ser Ala Gly Cys Thr Ala Thr Gly Gly Pro Glu Pro Pro Asp Ala Gly
            980                 985                 990

CCC AAG GCG CTT TAT GAT GTC GCA GAG GCA GAG GAG CGC TTC CCA GCG    3024
Pro Lys Ala Leu Tyr Asp Val Ala Glu Ala Glu Glu Arg Phe Pro Ala
        995                1000                1005

GCT GCC AGG CCG CGC TCG CCA TCG CCC ATC AGT ACG CTG AGC CAC CTG    3072
Ala Ala Arg Pro Arg Ser Pro Ser Pro Ile Ser Thr Leu Ser His Leu
    1010                1015                1020

GCA GGC TCG GCG GGC CGC ACA GAC GAC GAC GCG CCG TCG CTG CAC TCG    3120
Ala Gly Ser Ala Gly Arg Thr Asp Asp Asp Ala Pro Ser Leu His Ser
1025                1030                1035                1040

GAG ACC GCT GCA CGC AGC AGC TCA TCC CAG GGC TCG CTC ATG GAG CAG    3168
Glu Thr Ala Ala Arg Ser Ser Ser Ser Gln Gly Ser Leu Met Glu Gln
                1045                1050                1055

ATT AGC AGC GTG GTG ACG CGC TTC ACC GCC AAC ATC ACC GAG CTC AAC    3216
Ile Ser Ser Val Val Thr Arg Phe Thr Ala Asn Ile Thr Glu Leu Asn
            1060                1065                1070

TCC ATG ATG TTG TCC ACC GCG GCT GCG CCG GGG CCC CCT GGT ACC CCT    3264
Ser Met Met Leu Ser Thr Ala Ala Ala Pro Gly Pro Pro Gly Thr Pro
        1075                1080                1085

ATC TGC TCT TCC TAC CTG ATC CCC AAA GAG ATC CAG CTG CCC ACG ACC    3312
Ile Cys Ser Ser Tyr Leu Ile Pro Lys Glu Ile Gln Leu Pro Thr Thr
    1090                1095                1100

ATG ACG ACC TTC GCA GAG ATC CAG CCT CTG CCG GCC ATC GAG GTG ACC    3360
Met Thr Thr Phe Ala Glu Ile Gln Pro Leu Pro Ala Ile Glu Val Thr
1105                1110                1115                1120

GGA GGA GCT CAG CCG GCG ACA GGG CCA TCA CCT GCC CAA GAG ACG CCC    3408
Gly Gly Ala Gln Pro Ala Thr Gly Pro Ser Pro Ala Gln Glu Thr Pro
                1125                1130                1135

GCA GGA GCT GAA GCC GCC CCA GGA AAA CCG GAT CTG GAG GAG CTG GTG    3456
Ala Gly Ala Glu Ala Ala Pro Gly Lys Pro Asp Leu Glu Glu Leu Val
            1140                1145                1150

GCC CTC ACT CCA CCA TCG CCC TTC AGG GAC TCG GTG GAC TCG GGG AGC    3504
Ala Leu Thr Pro Pro Ser Pro Phe Arg Asp Ser Val Asp Ser Gly Ser
        1155                1160                1165

ACC ACC CCA AAC TCT CCA GTC TCC GAA TCG GCC CTC TGC ATC CCA TCC    3552
Thr Thr Pro Asn Ser Pro Val Ser Glu Ser Ala Leu Cys Ile Pro Ser
    1170                1175                1180

TCT CCC AAA TAT GAC ACT CTC ATC ATC AGA GAT TAC ACG                3591
Ser Pro Lys Tyr Asp Thr Leu Ile Ile Arg Asp Tyr Thr
1185                1190                1195
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1197 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Met Val Leu Leu Leu Ile Leu Ser Val Leu Leu Leu Lys Glu Asp Val
 1               5                  10                  15
Arg Gly Ser Ala Gln Ser Ser Glu Arg Arg Val Val Ala His Met Pro
            20                  25                  30
Gly Asp Ile Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro Thr
        35                  40                  45
Val Asp Lys Val His Glu Arg Lys Cys Gly Ala Val Arg Glu Gln Tyr
    50                  55                  60
Gly Ile Gln Arg Val Glu Ala Met Leu His Thr Leu Glu Arg Ile Asn
65                  70                  75                  80
Ser Asp Pro Thr Leu Leu Pro Asn Ile Thr Leu Gly Cys Glu Ile Arg
                85                  90                  95
Asp Ser Cys Trp His Ser Ala Val Ala Leu Glu Gln Ser Ile Glu Phe
            100                 105                 110
Ile Arg Asp Ser Leu Ile Ser Ser Glu Glu Glu Glu Gly Leu Val Arg
        115                 120                 125
Cys Val Asp Gly Ser Ser Ser Phe Arg Ser Lys Lys Pro Ile Val Gly
    130                 135                 140
Val Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val Gln Asn Leu
145                 150                 155                 160
Leu Gln Leu Phe Asn Ile Pro Gln Ile Ala Tyr Ser Ala Thr Ser Met
                165                 170                 175
Asp Leu Ser Asp Lys Thr Leu Phe Lys Tyr Phe Met Arg Val Val Pro
            180                 185                 190
Ser Asp Ala Gln Gln Ala Arg Ala Met Val Asp Ile Val Lys Arg Tyr
        195                 200                 205
Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn Tyr Gly Glu
    210                 215                 220
Ser Gly Met Glu Ala Phe Lys Asp Met Ser Ala Lys Glu Gly Ile Cys
225                 230                 235                 240
Ile Ala His Ser Tyr Lys Ile Tyr Ser Asn Ala Gly Glu Gln Ser Phe
                245                 250                 255
Asp Lys Leu Leu Lys Lys Leu Arg Ser His Leu Pro Lys Ala Arg Val
            260                 265                 270
Val Ala Cys Phe Cys Glu Gly Met Thr Val Arg Gly Leu Leu Met Ala
        275                 280                 285
Met Arg Arg Leu Gly Leu Ala Gly Glu Phe Leu Leu Leu Gly Ser Asp
    290                 295                 300
Gly Trp Ala Asp Arg Tyr Asp Val Thr Asp Gly Tyr Gln Arg Glu Ala
305                 310                 315                 320
Val Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro Asp Val Lys Trp Phe
                325                 330                 335
Asp Asp Tyr Tyr Leu Lys Leu Arg Pro Glu Thr Asn Leu Arg Asn Pro
            340                 345                 350
Trp Phe Gln Glu Phe Trp Gln His Arg Phe Gln Cys Arg Leu Glu Gly
        355                 360                 365
Phe Ala Gln Glu Asn Ser Lys Tyr Asn Lys Thr Cys Asn Ser Ser Leu
    370                 375                 380
Thr Leu Arg Thr His His Val Gln Asp Ser Lys Met Gly Phe Val Ile
```

385 390 395 400
Asn Ala Ile Tyr Ser Met Ala Tyr Gly Leu His Asn Met Gln Met Ser
405 410 415
Leu Cys Pro Gly Tyr Ala Gly Leu Cys Asp Ala Met Lys Pro Ile Asp
420 425 430
Gly Arg Lys Leu Leu Asp Ser Leu Met Lys Thr Asn Phe Thr Gly Val
435 440 445
Ser Gly Asp Met Ile Leu Phe Asp Glu Asn Gly Asp Ser Pro Gly Arg
450 455 460
Tyr Glu Ile Met Asn Phe Lys Glu Met Gly Lys Asp Tyr Phe Asp Tyr
465 470 475 480
Ile Asn Val Gly Ser Trp Asp Asn Gly Glu Leu Lys Met Asp Asp Asp
485 490 495
Glu Val Trp Ser Lys Lys Asn Asn Ile Ile Arg Ser Val Cys Ser Glu
500 505 510
Pro Cys Glu Lys Gly Gln Ile Lys Val Ile Arg Lys Gly Glu Val Ser
515 520 525
Cys Cys Trp Thr Cys Thr Pro Cys Lys Glu Asn Glu Tyr Val Phe Asp
530 535 540
Glu Tyr Thr Cys Lys Ala Cys Gln Leu Gly Ser Trp Pro Thr Asp Asp
545 550 555 560
Leu Thr Gly Cys Asp Leu Ile Pro Val Gln Tyr Leu Arg Trp Gly Asp
565 570 575
Pro Glu Pro Ile Ala Ala Val Val Phe Ala Cys Leu Gly Leu Leu Ala
580 585 590
Thr Leu Phe Val Thr Val Ile Phe Ile Ile Tyr Arg Asp Thr Pro Val
595 600 605
Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile Ile Leu Ala Gly Ile
610 615 620
Cys Leu Gly Tyr Leu Cys Thr Phe Cys Leu Ile Ala Lys Pro Lys Gln
625 630 635 640
Ile Tyr Cys Tyr Leu Gln Arg Ile Gly Ile Gly Leu Ser Pro Ala Met
645 650 655
Ser Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala Arg Ile Leu
660 665 670
Ala Gly Ser Lys Lys Lys Ile Cys Thr Lys Lys Pro Arg Phe Met Ser
675 680 685
Ala Cys Ala Gln Leu Val Ile Ala Phe Ile Leu Ile Cys Ile Gln Leu
690 695 700
Gly Ile Ile Val Ala Leu Phe Ile Met Glu Pro Pro Asp Ile Met His
705 710 715 720
Asp Tyr Pro Ser Ile Arg Glu Val Tyr Leu Ile Cys Asn Thr Thr Asn
725 730 735
Leu Gly Val Val Thr Pro Leu Gly Tyr Asn Gly Leu Leu Ile Leu Ser
740 745 750
Cys Thr Phe Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala Asn Phe Asn
755 760 765
Glu Ala Lys Tyr Ile Ala Phe Thr Met Tyr Thr Thr Cys Ile Ile Trp
770 775 780
Leu Ala Phe Val Pro Ile Tyr Phe Gly Ser Asn Tyr Lys Ile Ile Thr
785 790 795 800
Met Cys Phe Ser Val Ser Leu Ser Ala Thr Val Ala Leu Gly Cys Met
805 810 815
Phe Val Pro Lys Val Tyr Ile Ile Leu Ala Lys Pro Glu Arg Asn Val
820 825 830
Arg Ser Ala Phe Thr Thr Ser Thr Val Val Arg Met His Val Gly Asp
835 840 845
Gly Lys Ser Ser Ser Ala Ala Ser Arg Ser Ser Ser Leu Val Asn Leu
850 855 860
Trp Lys Arg Arg Gly Ser Ser Gly Glu Thr Leu Arg Tyr Lys Asp Arg
865 870 875 880

```
Arg Leu Ala Gln His Lys Ser Glu Ile Glu Cys Phe Thr Pro Lys Gly
                885                 890                 895
Ser Met Gly Asn Gly Gly Arg Ala Thr Met Ser Ser Ser Asn Gly Lys
            900                 905                 910
Ser Val Thr Trp Ala Gln Asn Glu Lys Ser Thr Arg Gly Gln His Leu
        915                 920                 925
Trp Gln Arg Leu Ser Val His Ile Asn Lys Lys Glu Asn Pro Asn Gln
    930                 935                 940
Thr Ala Val Ile Lys Pro Phe Pro Lys Ser Thr Glu Ser Arg Gly Gln
945                 950                 955                 960
Gly Ala Gly Ala Gly Gly Gly Ser Gly Pro Gly Ala Ala Gly Ala Gly
                965                 970                 975
Ser Ala Gly Cys Thr Ala Thr Gly Gly Pro Glu Pro Pro Asp Ala Gly
            980                 985                 990
Pro Lys Ala Leu Tyr Asp Val Ala Glu Ala Glu Glu Arg Phe Pro Ala
        995                1000                1005
Ala Ala Arg Pro Arg Ser Pro Ser Pro Ile Ser Thr Leu Ser His Leu
   1010                1015                1020
Ala Gly Ser Ala Gly Arg Thr Asp Asp Asp Ala Pro Ser Leu His Ser
025                1030                1035                1040
Glu Thr Ala Ala Arg Ser Ser Ser Ser Gln Gly Ser Leu Met Glu Gln
                1045                1050                1055
Ile Ser Ser Val Val Thr Arg Phe Thr Ala Asn Ile Thr Glu Leu Asn
            1060                1065                1070
Ser Met Met Leu Ser Thr Ala Ala Ala Pro Gly Pro Pro Gly Thr Pro
        1075                1080                1085
Ile Cys Ser Ser Tyr Leu Ile Pro Lys Glu Ile Gln Leu Pro Thr Thr
   1090                1095                1100
Met Thr Thr Phe Ala Glu Ile Gln Pro Leu Pro Ala Ile Glu Val Thr
105                1110                1115                1120
Gly Gly Ala Gln Pro Ala Thr Gly Pro Ser Pro Ala Gln Glu Thr Pro
                1125                1130                1135
Ala Gly Ala Glu Ala Ala Pro Gly Lys Pro Asp Leu Glu Glu Leu Val
            1140                1145                1150
Ala Leu Thr Pro Pro Ser Pro Phe Arg Asp Ser Val Asp Ser Gly Ser
        1155                1160                1165
Thr Thr Pro Asn Ser Pro Val Ser Glu Ser Ala Leu Cys Ile Pro Ser
   1170                1175                1180
Ser Pro Lys Tyr Asp Thr Leu Ile Ile Arg Asp Tyr Thr
185                1190                1195
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 100 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 3...98
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
AG GTA CAA AGA CAG GAG ACT GGC CCA GCA CAA GTC GGA AAT AGA GTG      47
   Val Gln Arg Gln Glu Thr Gly Pro Ala Gln Val Gly Asn Arg Val
    1               5                  10                  15
```

```
TTT CAC CCC CAA AGG GAG TAT GGG GAA TGG TGG GAG AGC AAC AAT GAG      95
Phe His Pro Gln Arg Glu Tyr Gly Glu Trp Trp Glu Ser Asn Asn Glu
                20                  25                  30

CAG GT                                                              100
Gln
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Val Gln Arg Gln Glu Thr Gly Pro Ala Gln Val Gly Asn Arg Val Phe
 1               5                  10                  15
His Pro Gln Arg Glu Tyr Gly Glu Trp Trp Glu Ser Asn Asn Glu Gln
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3918 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence
(B) LOCATION: 268...3780
(D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
TCAGACATCT CTTGAGAAAC TCAGCCTAGT TTATCCAGCT GTTTTGTCCA CATATGTTCA      60
TAAAATCTAG TTATGAATAT TCTGATCGTA TGGATGCACC TAACTCGACA GTTTCTGGGG     120
ACTTTGCTAT CTTTTAGAGA ATATATCTAA ATTGCTGACT AATTCCTCAA TGCGCCACTC     180
AACGCAGGGC ATCATTTGGC TGTACCTATT GGAGCTCTGA GTTTCCTCCA TCATGATATC     240
TTTATTAGCT TGAATTCCTT TCCTAAA ATG GTC CTT CTG TTG ATC CTG TCA GTC     294
                              Met Val Leu Leu Leu Ile Leu Ser Val
                               1               5

CTA CTT CTG AAA GAA GAT GTA CGA GGG AGT GCA CAG TCC AGT GAG AGG      342
Leu Leu Leu Lys Glu Asp Val Arg Gly Ser Ala Gln Ser Ser Glu Arg
10                  15                  20                  25

AGG GTG GTG GCT CAC ATG CCA GGT GAC ATC ATT ATT GGA GCT CTC TTC      390
Arg Val Val Ala His Met Pro Gly Asp Ile Ile Ile Gly Ala Leu Phe
                30                  35                  40

TCT GTC CAC CAC CAA CCA ACT GTG GAC AAA GTT CAT GAG AGG AAG TGT      438
Ser Val His His Gln Pro Thr Val Asp Lys Val His Glu Arg Lys Cys
            45                  50                  55
```

```
GGT GCA GTT CGT GAG CAA TAT GGG ATT CAG AGA GTG GAG GCC ATG CTG      486
Gly Ala Val Arg Glu Gln Tyr Gly Ile Gln Arg Val Glu Ala Met Leu
        60                  65                  70

CAT ACC TTG GAA AGG ATC AAC TCA GAT CCC ACA CTC TTG CCC AAC ATC      534
His Thr Leu Glu Arg Ile Asn Ser Asp Pro Thr Leu Leu Pro Asn Ile
    75                  80                  85

ACA CTT GGC TGT GAG ATC AGA GAT TCC TGC TGG CAT TCC GCT GTG GCC      582
Thr Leu Gly Cys Glu Ile Arg Asp Ser Cys Trp His Ser Ala Val Ala
90                  95                  100                 105

CTA GAA CAA AGC ATT GAG TTT ATA AGG GAT TCC CTC ATC TCT TCA GAA      630
Leu Glu Gln Ser Ile Glu Phe Ile Arg Asp Ser Leu Ile Ser Ser Glu
                110                 115                 120

GAG GAG GAA GGT TTG GTA CGC TGT GTA GAT GGA TCT TCT TCC TTC CGC      678
Glu Glu Glu Gly Leu Val Arg Cys Val Asp Gly Ser Ser Ser Phe Arg
            125                 130                 135

TCC AAG AAA CCC ATA GTG GGA GTC ATT GGG CCT GGC TCG AGT TCT GTG      726
Ser Lys Lys Pro Ile Val Gly Val Ile Gly Pro Gly Ser Ser Ser Val
        140                 145                 150

GCC ATT CAA GTC CAG AAC TTA CTC CAG CTT TTC AAC ATA CCA CAG ATT      774
Ala Ile Gln Val Gln Asn Leu Leu Gln Leu Phe Asn Ile Pro Gln Ile
    155                 160                 165

GCT TAC TCT GCA ACC AGC ATG GAT TTG AGT GAC AAG ACT CTA TTC AAG      822
Ala Tyr Ser Ala Thr Ser Met Asp Leu Ser Asp Lys Thr Leu Phe Lys
170                 175                 180                 185

TAC TTC ATG AGG GTT GTA CCT TCC GAT GCC CAG CAA GCT CGA GCC ATG      870
Tyr Phe Met Arg Val Val Pro Ser Asp Ala Gln Gln Ala Arg Ala Met
                190                 195                 200

GTA GAC ATA GTG AAG AGA TAC AAC TGG ACT TAT GTC TCA GCT GTG CAC      918
Val Asp Ile Val Lys Arg Tyr Asn Trp Thr Tyr Val Ser Ala Val His
            205                 210                 215

ACA GAA GGC AAC TAT GGA GAA AGT GGG ATG GAG GCT TTC AAA GAT ATG      966
Thr Glu Gly Asn Tyr Gly Glu Ser Gly Met Glu Ala Phe Lys Asp Met
        220                 225                 230

TCA GCG AAG GAA GGG ATT TGC ATC GCC CAC TCT TAC AAA ATT TAC AGC     1014
Ser Ala Lys Glu Gly Ile Cys Ile Ala His Ser Tyr Lys Ile Tyr Ser
    235                 240                 245

AAT GCT GGG GAA CAG AGC TTT GAC AAG CTG TTG AAA AAA CTC AGA AGT     1062
Asn Ala Gly Glu Gln Ser Phe Asp Lys Leu Leu Lys Lys Leu Arg Ser
250                 255                 260                 265

CAT TTA CCT AAA GCC CGG GTG GTA GCT TGC TTC TGT GAA GGC ATG ACA     1110
His Leu Pro Lys Ala Arg Val Val Ala Cys Phe Cys Glu Gly Met Thr
                270                 275                 280

GTT CGA GGT CTG CTC ATG GCC ATG AGA CGC TTA GGT CTA GCA GGG GAG     1158
Val Arg Gly Leu Leu Met Ala Met Arg Arg Leu Gly Leu Ala Gly Glu
            285                 290                 295

TTT CTA CTT CTG GGC AGT GAT GGC TGG GCC GAC AGG TAT GAC GTG ACA     1206
```

```
Phe Leu Leu Leu Gly Ser Asp Gly Trp Ala Asp Arg Tyr Asp Val Thr
        300                 305                 310

GAT GGA TAT CAG CGA GAA GCT GTC GGT GGA ATT ACA ATC AAG CTT CAG    1254
Asp Gly Tyr Gln Arg Glu Ala Val Gly Gly Ile Thr Ile Lys Leu Gln
    315                 320                 325

TCT CCT GAT GTC AAG TGG TTT GAT GAT TAT TAT CTG AAA CTC CGG CCA    1302
Ser Pro Asp Val Lys Trp Phe Asp Asp Tyr Tyr Leu Lys Leu Arg Pro
330                 335                 340                 345

GAA ACA AAC CTC AGA AAC CCT TGG TTT CAA GAA TTT TGG CAG CAT CGT    1350
Glu Thr Asn Leu Arg Asn Pro Trp Phe Gln Glu Phe Trp Gln His Arg
                350                 355                 360

TTT CAG TGC CGG CTG GAA GGG TTT GCA CAG GAG AAC AGC AAG TAC AAC    1398
Phe Gln Cys Arg Leu Glu Gly Phe Ala Gln Glu Asn Ser Lys Tyr Asn
            365                 370                 375

AAG ACT TGC AAC AGT TCT CTG ACT CTG AGA ACA CAT CAT GTT CAA GAT    1446
Lys Thr Cys Asn Ser Ser Leu Thr Leu Arg Thr His His Val Gln Asp
        380                 385                 390

TCC AAA ATG GGA TTT GTG ATA AAT GCA ATC TAT TCT ATG GCT TAT GGA    1494
Ser Lys Met Gly Phe Val Ile Asn Ala Ile Tyr Ser Met Ala Tyr Gly
    395                 400                 405

CTC CAC AAT ATG CAG ATG TCC CTG TGT CCA GGC TAT GCA GGC CTC TGT    1542
Leu His Asn Met Gln Met Ser Leu Cys Pro Gly Tyr Ala Gly Leu Cys
410                 415                 420                 425

GAT GCG ATG AAG CCA ATT GAT GGA CGG AAA CTT TTG GAT TCC CTG ATG    1590
Asp Ala Met Lys Pro Ile Asp Gly Arg Lys Leu Leu Asp Ser Leu Met
                430                 435                 440

AAA ACC AAT TTT ACT GGA GTT TCT GGA GAT ATG ATC CTG TTT GAT GAA    1638
Lys Thr Asn Phe Thr Gly Val Ser Gly Asp Met Ile Leu Phe Asp Glu
            445                 450                 455

AAT GGA GAC TCT CCA GGA AGG TAT GAA ATA ATG AAT TTC AAG GAA ATG    1686
Asn Gly Asp Ser Pro Gly Arg Tyr Glu Ile Met Asn Phe Lys Glu Met
        460                 465                 470

GGA AAA GAT TAT TTT GAT TAC ATC AAT GTT GGA AGT TGG GAC AAT GGG    1734
Gly Lys Asp Tyr Phe Asp Tyr Ile Asn Val Gly Ser Trp Asp Asn Gly
    475                 480                 485

GAA TTA AAA ATG GAT GAT GAC GAA GTG TGG TCC AAG AAA AAT AAC ATC    1782
Glu Leu Lys Met Asp Asp Asp Glu Val Trp Ser Lys Lys Asn Asn Ile
490                 495                 500                 505

ATC AGA TCT GTG TGC AGT GAA CCA TGT GAG AAA GGC CAA ATA AAG GTG    1830
Ile Arg Ser Val Cys Ser Glu Pro Cys Glu Lys Gly Gln Ile Lys Val
                510                 515                 520

ATC CGG AAG GGA GAA GTC AGC TGT TGT TGG ACC TGC ACA CCT TGT AAA    1878
Ile Arg Lys Gly Glu Val Ser Cys Cys Trp Thr Cys Thr Pro Cys Lys
            525                 530                 535

GAG AAT GAG TAT GTG TTT GAT GAG TAC ACC TGC AAG GCA TGC CAA CTG    1926
Glu Asn Glu Tyr Val Phe Asp Glu Tyr Thr Cys Lys Ala Cys Gln Leu
```

```
                540                 545                 550

GGC TCC TGG CCC ACT GAC GAC CTG ACA GGC TGT GAC TTG ATC CCA GTC    1974
Gly Ser Trp Pro Thr Asp Asp Leu Thr Gly Cys Asp Leu Ile Pro Val
    555                 560                 565

CAG TAT CTT CGC TGG GGT GAC CCT GAG CCG ATT GCA GCT GTG GTG TTT    2022
Gln Tyr Leu Arg Trp Gly Asp Pro Glu Pro Ile Ala Ala Val Val Phe
570                 575                 580                 585

GCC TGC CTC GGT CTG CTA GCC ACC TTA TTC GTT ACT GTA ATC TTC ATC    2070
Ala Cys Leu Gly Leu Leu Ala Thr Leu Phe Val Thr Val Ile Phe Ile
                590                 595                 600

ATT TAT CGG GAC ACT CCG GTG GTC AAG TCC TCC AGC AGG GAG CTC TGC    2118
Ile Tyr Arg Asp Thr Pro Val Val Lys Ser Ser Ser Arg Glu Leu Cys
            605                 610                 615

TAC ATT ATC CTT GCT GGC ATC TGC CTG GGA TAC TTA TGT ACC TTT TGC    2166
Tyr Ile Ile Leu Ala Gly Ile Cys Leu Gly Tyr Leu Cys Thr Phe Cys
        620                 625                 630

CTC ATT GCA AAG CCC AAG CAG ATT TAC TGC TAT CTT CAG AGA ATT GGC    2214
Leu Ile Ala Lys Pro Lys Gln Ile Tyr Cys Tyr Leu Gln Arg Ile Gly
    635                 640                 645

ATT GGA CTC TCT CCA GCC ATG AGC TAC TCA GCC CTT GTA ACC AAG ACC    2262
Ile Gly Leu Ser Pro Ala Met Ser Tyr Ser Ala Leu Val Thr Lys Thr
650                 655                 660                 665

AAC CGT ATT GCA AGG ATT CTA GCT GGA AGC AAG AAG AAA ATC TGT ACC    2310
Asn Arg Ile Ala Arg Ile Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr
                670                 675                 680

AAG AAG CCC AGA TTC ATG AGC GCC TGT GCT CAG TTA GTG ATC GCT TTC    2358
Lys Lys Pro Arg Phe Met Ser Ala Cys Ala Gln Leu Val Ile Ala Phe
            685                 690                 695

ATT CTC ATC TGT ATT CAG CTG GGC ATT ATC GTG GCC CTC TTT ATC ATG    2406
Ile Leu Ile Cys Ile Gln Leu Gly Ile Ile Val Ala Leu Phe Ile Met
        700                 705                 710

GAG CCT CCG GAT ATA ATG CAT GAC TAT CCA AGC ATC CGA GAA GTC TAC    2454
Glu Pro Pro Asp Ile Met His Asp Tyr Pro Ser Ile Arg Glu Val Tyr
    715                 720                 725

TTG ATT TGT AAC ACC ACC AAC CTA GGG GTT GTT ACT CCT CTT GGA TAC    2502
Leu Ile Cys Asn Thr Thr Asn Leu Gly Val Val Thr Pro Leu Gly Tyr
730                 735                 740                 745

AAT GGA TTA TTG ATT TTG AGT TGC ACG TTC TAT GCG TTT AAG ACC AGA    2550
Asn Gly Leu Leu Ile Leu Ser Cys Thr Phe Tyr Ala Phe Lys Thr Arg
                750                 755                 760

AAT GTT CCA GCC AAC TTT AAT GAG GCC AAA TAT ATT GCT TTC ACC ATG    2598
Asn Val Pro Ala Asn Phe Asn Glu Ala Lys Tyr Ile Ala Phe Thr Met
            765                 770                 775

TAC ACA ACC TGC ATC ATA TGG CTG GCC TTC GTG CCT ATC TAC TTT GGC    2646
Tyr Thr Thr Cys Ile Ile Trp Leu Ala Phe Val Pro Ile Tyr Phe Gly
        780                 785                 790
```

```
AGC AAC TAC AAA ATC ATC ACC ATG TGT TTT TCA GTC AGC CTC AGT GCC    2694
Ser Asn Tyr Lys Ile Ile Thr Met Cys Phe Ser Val Ser Leu Ser Ala
    795                 800                 805

ACA GTG GCC CTG GGT TGC ATG TTT GTC CCG AAG GTG TAC ATC ATC CTA    2742
Thr Val Ala Leu Gly Cys Met Phe Val Pro Lys Val Tyr Ile Ile Leu
810                 815                 820                 825

GCC AAA CCG GAG AGA AAT GTG CGC AGC GCC TTC ACA ACC TCT ACA GTG    2790
Ala Lys Pro Glu Arg Asn Val Arg Ser Ala Phe Thr Thr Ser Thr Val
                830                 835                 840

GTA CGC ATG CAT GTA GGA GAC GGC AAA TCA TCG TCC GCT GCC AGC AGA    2838
Val Arg Met His Val Gly Asp Gly Lys Ser Ser Ser Ala Ala Ser Arg
            845                 850                 855

TCC AGC AGC CTA GTC AAC CTG TGG AAG AGG AGG GGC TCC TCT GGG GAA    2886
Ser Ser Ser Leu Val Asn Leu Trp Lys Arg Arg Gly Ser Ser Gly Glu
        860                 865                 870

ACC CTA AGC TCC AAC GGA AAA TCT GTG ACT TGG GCC CAG AAT GAG AAG    2934
Thr Leu Ser Ser Asn Gly Lys Ser Val Thr Trp Ala Gln Asn Glu Lys
    875                 880                 885

AGT ACC CGG GGG CAA CAT TTG TGG CAG CGA CTG TCT GTC CAC ATC AAC    2982
Ser Thr Arg Gly Gln His Leu Trp Gln Arg Leu Ser Val His Ile Asn
890                 895                 900                 905

AAG AAG GAG AAC CCC AAC CAA ACG GCG GTC ATC AAA CCC TTT CCC AAG    3030
Lys Lys Glu Asn Pro Asn Gln Thr Ala Val Ile Lys Pro Phe Pro Lys
                910                 915                 920

AGC ACA GAG AAC CGG GGG CCT GGT GCA GCG GCA GGT GGT GGC TCG GGT    3078
Ser Thr Glu Asn Arg Gly Pro Gly Ala Ala Ala Gly Gly Gly Ser Gly
            925                 930                 935

CCC GGT GTA GCT GGT GCT GGT AAC GCA GGA TGC ACA GCA ACA GGT GGC    3126
Pro Gly Val Ala Gly Ala Gly Asn Ala Gly Cys Thr Ala Thr Gly Gly
        940                 945                 950

CCG GAG CCA CCG GAC GCC GGT CCC AAA GCG CTT TAT GAT GTG GCG GAG    3174
Pro Glu Pro Pro Asp Ala Gly Pro Lys Ala Leu Tyr Asp Val Ala Glu
    955                 960                 965

GCA GAG GAG AGC TTC CCC GCG GCT GCC AGG CCG CGC TCG CCA TCG CCC    3222
Ala Glu Glu Ser Phe Pro Ala Ala Ala Arg Pro Arg Ser Pro Ser Pro
970                 975                 980                 985

ATC AGC ACG TTG AGC CAC CTC GCA GGC TCG GCG GGC CGC ACA GAT GAT    3270
Ile Ser Thr Leu Ser His Leu Ala Gly Ser Ala Gly Arg Thr Asp Asp
                990                 995                1000

GAT GCG CCG TCG CTG CAC TCT GAG ACA GCT GCT CGC AGC AGC TCG TCC    3318
Asp Ala Pro Ser Leu His Ser Glu Thr Ala Ala Arg Ser Ser Ser Ser
            1005                1010                1015

CAG GGC TCG CTC ATG GAG CAG ATC AGC AGC GTA GTG ACG CGC TTC ACA    3366
Gln Gly Ser Leu Met Glu Gln Ile Ser Ser Val Val Thr Arg Phe Thr
        1020                1025                1030
```

```
GCC AAC ATC TCC GAG CTC AAC TCC ATG ATG TTG TCC ACC GCG GCC ACG    3414
Ala Asn Ile Ser Glu Leu Asn Ser Met Met Leu Ser Thr Ala Ala Thr
   1035                1040                1045

CCG GGG CCC CCT GGT ACT CCA ATC TGC TCC TCC TAC CTG ATC CCC AAA    3462
Pro Gly Pro Pro Gly Thr Pro Ile Cys Ser Ser Tyr Leu Ile Pro Lys
1050                1055                1060                1065

GAG ATT CAG CTG CCC ACG ACC ATG ACG ACC TTC GCA GAG ATC CAG CCG    3510
Glu Ile Gln Leu Pro Thr Thr Met Thr Thr Phe Ala Glu Ile Gln Pro
                1070                1075                1080

CTG CCG GCC ATC GAG GTG ACC GGA GGA GCT CAG GGG GCG ACA GGC GTA    3558
Leu Pro Ala Ile Glu Val Thr Gly Gly Ala Gln Gly Ala Thr Gly Val
           1085                1090                1095

TCA CCT GCC CAG GAG ACG CCC ACA GGA GCT GAA TCC GCC CCG GGC AAA    3606
Ser Pro Ala Gln Glu Thr Pro Thr Gly Ala Glu Ser Ala Pro Gly Lys
       1100                1105                1110

CCA GAT CTG GAG GAG CTT GTG GCC CTC ACT CCA CCA TCG CCC TTC AGG    3654
Pro Asp Leu Glu Glu Leu Val Ala Leu Thr Pro Pro Ser Pro Phe Arg
   1115                1120                1125

GAC TCG GTG GAC TCG GGG AGC ACC ACC CCA AAC TCT CCA GTC TCA GAA    3702
Asp Ser Val Asp Ser Gly Ser Thr Thr Pro Asn Ser Pro Val Ser Glu
1130                1135                1140                1145

TCG GCC CTC TGC ATC CCA TCC TCT CCC AAA TAT GAC ACT CTC ATC ATC    3750
Ser Ala Leu Cys Ile Pro Ser Ser Pro Lys Tyr Asp Thr Leu Ile Ile
                1150                1155                1160

AGA GAT TAC ACG CAG AGT TCT TCA TCG TTG TGAGCCACTG GAAACTTCCC TGA  3803
Arg Asp Tyr Thr Gln Ser Ser Ser Ser Leu
           1165                1170

GATAGGCCTG TGCAGCAGAC CTCGGTGTTC CAGATGGGCG CAGCGACACC CATAGTCACC   3863
CCAACAGAGA ACCCGAGGTG AAAAATGCCA AGGACACACA CGGTGGAGAC ATGAG        3918
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1171 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Met Val Leu Leu Leu Ile Leu Ser Val Leu Leu Leu Lys Glu Asp Val
 1               5                  10                  15
Arg Gly Ser Ala Gln Ser Ser Glu Arg Arg Val Val Ala His Met Pro
            20                  25                  30
Gly Asp Ile Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro Thr
        35                  40                  45
Val Asp Lys Val His Glu Arg Lys Cys Gly Ala Val Arg Glu Gln Tyr
    50                  55                  60
Gly Ile Gln Arg Val Glu Ala Met Leu His Thr Leu Glu Arg Ile Asn
```

65 70 75 80
Ser Asp Pro Thr Leu Leu Pro Asn Ile Thr Leu Gly Cys Glu Ile Arg
85 90 95
Asp Ser Cys Trp His Ser Ala Val Ala Leu Glu Gln Ser Ile Glu Phe
100 105 110
Ile Arg Asp Ser Leu Ile Ser Ser Glu Glu Glu Glu Gly Leu Val Arg
115 120 125
Cys Val Asp Gly Ser Ser Ser Phe Arg Ser Lys Lys Pro Ile Val Gly
130 135 140
Val Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val Gln Asn Leu
145 150 155 160
Leu Gln Leu Phe Asn Ile Pro Gln Ile Ala Tyr Ser Ala Thr Ser Met
165 170 175
Asp Leu Ser Asp Lys Thr Leu Phe Lys Tyr Phe Met Arg Val Val Pro
180 185 190
Ser Asp Ala Gln Gln Ala Arg Ala Met Val Asp Ile Val Lys Arg Tyr
195 200 205
Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn Tyr Gly Glu
210 215 220
Ser Gly Met Glu Ala Phe Lys Asp Met Ser Ala Lys Glu Gly Ile Cys
225 230 235 240
Ile Ala His Ser Tyr Lys Ile Tyr Ser Asn Ala Gly Glu Gln Ser Phe
245 250 255
Asp Lys Leu Leu Lys Lys Leu Arg Ser His Leu Pro Lys Ala Arg Val
260 265 270
Val Ala Cys Phe Cys Glu Gly Met Thr Val Arg Gly Leu Leu Met Ala
275 280 285
Met Arg Arg Leu Gly Leu Ala Gly Glu Phe Leu Leu Leu Gly Ser Asp
290 295 300
Gly Trp Ala Asp Arg Tyr Asp Val Thr Asp Gly Tyr Gln Arg Glu Ala
305 310 315 320
Val Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro Asp Val Lys Trp Phe
325 330 335
Asp Asp Tyr Tyr Leu Lys Leu Arg Pro Glu Thr Asn Leu Arg Asn Pro
340 345 350
Trp Phe Gln Glu Phe Trp Gln His Arg Phe Gln Cys Arg Leu Glu Gly
355 360 365
Phe Ala Gln Glu Asn Ser Lys Tyr Asn Lys Thr Cys Asn Ser Ser Leu
370 375 380
Thr Leu Arg Thr His His Val Gln Asp Ser Lys Met Gly Phe Val Ile
385 390 395 400
Asn Ala Ile Tyr Ser Met Ala Tyr Gly Leu His Asn Met Gln Met Ser
405 410 415
Leu Cys Pro Gly Tyr Ala Gly Leu Cys Asp Ala Met Lys Pro Ile Asp
420 425 430
Gly Arg Lys Leu Leu Asp Ser Leu Met Lys Thr Asn Phe Thr Gly Val
435 440 445
Ser Gly Asp Met Ile Leu Phe Asp Glu Asn Gly Asp Ser Pro Gly Arg
450 455 460
Tyr Glu Ile Met Asn Phe Lys Glu Met Gly Lys Asp Tyr Phe Asp Tyr
465 470 475 480
Ile Asn Val Gly Ser Trp Asp Asn Gly Glu Leu Lys Met Asp Asp Asp
485 490 495
Glu Val Trp Ser Lys Lys Asn Asn Ile Ile Arg Ser Val Cys Ser Glu
500 505 510
Pro Cys Glu Lys Gly Gln Ile Lys Val Ile Arg Lys Gly Glu Val Ser
515 520 525
Cys Cys Trp Thr Cys Thr Pro Cys Lys Glu Asn Glu Tyr Val Phe Asp
530 535 540
Glu Tyr Thr Cys Lys Ala Cys Gln Leu Gly Ser Trp Pro Thr Asp Asp
545 550 555 560

Leu Thr Gly Cys Asp Leu Ile Pro Val Gln Tyr Leu Arg Trp Gly Asp
565 570 575
Pro Glu Pro Ile Ala Ala Val Val Phe Ala Cys Leu Gly Leu Leu Ala
580 585 590
Thr Leu Phe Val Thr Val Ile Phe Ile Ile Tyr Arg Asp Thr Pro Val
595 600 605
Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile Ile Leu Ala Gly Ile
610 615 620
Cys Leu Gly Tyr Leu Cys Thr Phe Cys Leu Ile Ala Lys Pro Lys Gln
625 630 635 640
Ile Tyr Cys Tyr Leu Gln Arg Ile Gly Ile Gly Leu Ser Pro Ala Met
645 650 655
Ser Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala Arg Ile Leu
660 665 670
Ala Gly Ser Lys Lys Lys Ile Cys Thr Lys Lys Pro Arg Phe Met Ser
675 680 685
Ala Cys Ala Gln Leu Val Ile Ala Phe Ile Leu Ile Cys Ile Gln Leu
690 695 700
Gly Ile Ile Val Ala Leu Phe Ile Met Glu Pro Pro Asp Ile Met His
705 710 715 720
Asp Tyr Pro Ser Ile Arg Glu Val Tyr Leu Ile Cys Asn Thr Thr Asn
725 730 735
Leu Gly Val Val Thr Pro Leu Gly Tyr Asn Gly Leu Leu Ile Leu Ser
740 745 750
Cys Thr Phe Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala Asn Phe Asn
755 760 765
Glu Ala Lys Tyr Ile Ala Phe Thr Met Tyr Thr Thr Cys Ile Ile Trp
770 775 780
Leu Ala Phe Val Pro Ile Tyr Phe Gly Ser Asn Tyr Lys Ile Ile Thr
785 790 795 800
Met Cys Phe Ser Val Ser Leu Ser Ala Thr Val Ala Leu Gly Cys Met
805 810 815
Phe Val Pro Lys Val Tyr Ile Ile Leu Ala Lys Pro Glu Arg Asn Val
820 825 830
Arg Ser Ala Phe Thr Thr Ser Thr Val Val Arg Met His Val Gly Asp
835 840 845
Gly Lys Ser Ser Ser Ala Ala Ser Arg Ser Ser Ser Leu Val Asn Leu
850 855 860
Trp Lys Arg Arg Gly Ser Ser Gly Glu Thr Leu Ser Ser Asn Gly Lys
865 870 875 880
Ser Val Thr Trp Ala Gln Asn Glu Lys Ser Thr Arg Gly Gln His Leu
885 890 895
Trp Gln Arg Leu Ser Val His Ile Asn Lys Lys Glu Asn Pro Asn Gln
900 905 910
Thr Ala Val Ile Lys Pro Phe Pro Lys Ser Thr Glu Asn Arg Gly Pro
915 920 925
Gly Ala Ala Ala Gly Gly Gly Ser Gly Pro Gly Val Ala Gly Ala Gly
930 935 940
Asn Ala Gly Cys Thr Ala Thr Gly Gly Pro Glu Pro Pro Asp Ala Gly
945 950 955 960
Pro Lys Ala Leu Tyr Asp Val Ala Glu Ala Glu Glu Ser Phe Pro Ala
965 970 975
Ala Ala Arg Pro Arg Ser Pro Ser Pro Ile Ser Thr Leu Ser His Leu
980 985 990
Ala Gly Ser Ala Gly Arg Thr Asp Asp Asp Ala Pro Ser Leu His Ser
995 1000 1005
Glu Thr Ala Ala Arg Ser Ser Ser Ser Gln Gly Ser Leu Met Glu Gln
1010 1015 1020
Ile Ser Ser Val Val Thr Arg Phe Thr Ala Asn Ile Ser Glu Leu Asn
1025 1030 1035 1040
Ser Met Met Leu Ser Thr Ala Ala Thr Pro Gly Pro Pro Gly Thr Pro

```
                1045                1050                1055
Ile Cys Ser Ser Tyr Leu Ile Pro Lys Glu Ile Gln Leu Pro Thr Thr
            1060                1065                1070
Met Thr Thr Phe Ala Glu Ile Gln Pro Leu Pro Ala Ile Glu Val Thr
        1075                1080                1085
Gly Gly Ala Gln Gly Ala Thr Gly Val Ser Pro Ala Gln Glu Thr Pro
    1090                1095                1100
Thr Gly Ala Glu Ser Ala Pro Gly Lys Pro Asp Leu Glu Glu Leu Val
105                 1110                1115                1120
Ala Leu Thr Pro Pro Ser Pro Phe Arg Asp Ser Val Asp Ser Gly Ser
                1125                1130                1135
Thr Thr Pro Asn Ser Pro Val Ser Glu Ser Ala Leu Cys Ile Pro Ser
            1140                1145                1150
Ser Pro Lys Tyr Asp Thr Leu Ile Ile Arg Asp Tyr Thr Gln Ser Ser
        1155                1160                1165
Ser Ser Leu
    1170
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1212 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Met Val Leu Leu Leu Ile Leu Ser Val Leu Leu Leu Lys Glu Asp Val
 1               5                  10                  15
Arg Gly Ser Ala Gln Ser Ser Glu Arg Arg Val Val Ala His Met Pro
            20                  25                  30
Gly Asp Ile Ile Ile Gly Ala Leu Phe Ser Val His His Gln Pro Thr
        35                  40                  45
Val Asp Lys Val His Glu Arg Lys Cys Gly Ala Val Arg Glu Gln Tyr
    50                  55                  60
Gly Ile Gln Arg Val Glu Ala Met Leu His Thr Leu Glu Arg Ile Asn
65                  70                  75                  80
Ser Asp Pro Thr Leu Leu Pro Asn Ile Thr Leu Gly Cys Glu Ile Arg
                85                  90                  95
Asp Ser Cys Trp His Ser Ala Val Ala Leu Glu Gln Ser Ile Glu Phe
            100                 105                 110
Ile Arg Asp Ser Leu Ile Ser Ser Glu Glu Glu Glu Gly Leu Val Arg
        115                 120                 125
Cys Val Asp Gly Ser Ser Ser Ser Phe Arg Ser Lys Lys Pro Ile Val
    130                 135                 140
Gly Val Ile Gly Pro Gly Ser Ser Ser Val Ala Ile Gln Val Gln Asn
145                 150                 155                 160
Leu Leu Gln Leu Phe Asn Ile Pro Gln Ile Ala Tyr Ser Ala Thr Ser
                165                 170                 175
Met Asp Leu Ser Asp Lys Thr Leu Phe Lys Tyr Phe Met Arg Val Val
            180                 185                 190
Pro Ser Asp Ala Gln Gln Ala Arg Ala Met Val Asp Ile Val Lys Arg
        195                 200                 205
Tyr Asn Trp Thr Tyr Val Ser Ala Val His Thr Glu Gly Asn Tyr Gly
    210                 215                 220
Glu Ser Gly Met Glu Ala Phe Lys Asp Met Ser Ala Lys Glu Gly Ile
225                 230                 235                 240
Cys Ile Ala His Ser Tyr Lys Ile Tyr Ser Asn Ala Gly Glu Gln Ser
                245                 250                 255
Phe Asp Lys Leu Leu Lys Lys Leu Thr Ser His Leu Pro Lys Ala Arg
```

260 265 270
Val Val Ala Cys Phe Cys Glu Gly Met Thr Val Arg Gly Leu Leu Met
275 280 285
Ala Met Arg Arg Leu Gly Leu Ala Gly Glu Phe Leu Leu Leu Gly Ser
290 295 300
Asp Gly Trp Ala Asp Arg Tyr Asp Val Thr Asp Gly Tyr Gln Arg Glu
305 310 315 320
Ala Val Gly Gly Ile Thr Ile Lys Leu Gln Ser Pro Asp Val Lys Trp
325 330 335
Phe Asp Asp Tyr Tyr Leu Lys Leu Arg Pro Glu Thr Asn His Arg Asn
340 345 350
Pro Trp Phe Gln Glu Phe Trp Gln His Arg Phe Gln Cys Arg Leu Glu
355 360 365
Gly Phe Pro Gln Glu Asn Ser Lys Tyr Asn Lys Thr Cys Asn Ser Ser
370 375 380
Leu Thr Leu Lys Thr His His Val Gln Asp Ser Lys Met Gly Phe Val
385 390 395 400
Ile Asn Ala Ile Tyr Ser Met Ala Tyr Gly Leu His Asn Met Gln Met
405 410 415
Ser Leu Cys Pro Gly Tyr Ala Gly Leu Cys Asp Ala Met Lys Pro Ile
420 425 430
Asp Gly Arg Lys Leu Leu Glu Ser Leu Met Lys Thr Asn Phe Thr Gly
435 440 445
Val Ser Gly Asp Thr Ile Leu Phe Asp Glu Asn Gly Asp Ser Pro Gly
450 455 460
Arg Tyr Glu Ile Met Asn Phe Lys Glu Met Gly Lys Asp Tyr Phe Asp
465 470 475 480
Tyr Ile Asn Val Gly Ser Trp Asp Asn Gly Glu Leu Lys Met Asp Asp
485 490 495
Asp Glu Val Trp Ser Lys Lys Ser Asn Ile Ile Arg Ser Val Cys Ser
500 505 510
Glu Pro Cys Glu Lys Gly Gln Ile Lys Val Ile Arg Lys Gly Glu Val
515 520 525
Ser Cys Cys Trp Thr Cys Thr Pro Cys Lys Glu Asn Glu Tyr Val Phe
530 535 540
Asp Glu Tyr Thr Cys Lys Ala Cys Gln Leu Gly Ser Trp Pro Thr Asp
545 550 555 560
Asp Leu Thr Gly Cys Asp Leu Ile Pro Val Gln Tyr Leu Arg Trp Gly
565 570 575
Asp Pro Glu Pro Ile Ala Ala Val Val Phe Ala Cys Leu Gly Leu Leu
580 585 590
Ala Thr Leu Phe Val Thr Val Val Phe Ile Ile Tyr Arg Asp Thr Pro
595 600 605
Val Val Lys Ser Ser Ser Arg Glu Leu Cys Tyr Ile Ile Leu Ala Gly
610 615 620
Ile Cys Leu Gly Tyr Leu Cys Thr Phe Cys Leu Ile Ala Lys Pro Lys
625 630 635 640
Gln Ile Tyr Cys Tyr Leu Gln Arg Ile Gly Ile Gly Leu Ser Pro Ala
645 650 655
Met Ser Tyr Ser Ala Leu Val Thr Lys Thr Asn Arg Ile Ala Arg Ile
660 665 670
Leu Ala Gly Ser Lys Lys Lys Ile Cys Thr Lys Lys Pro Arg Phe Met
675 680 685
Ser Ala Cys Ala Gln Leu Val Ile Ala Phe Ile Leu Ile Cys Ile Gln
690 695 700
Leu Gly Ile Ile Val Ala Leu Phe Ile Met Glu Pro Pro Asp Ile Met
705 710 715 720
His Asp Tyr Pro Ser Ile Arg Glu Val Tyr Leu Ile Cys Asn Thr Thr
725 730 735
Asn Leu Gly Val Val Thr Pro Leu Gly Tyr Asn Gly Leu Leu Ile Leu
740 745 750

```
Ser Cys Thr Phe Tyr Ala Phe Lys Thr Arg Asn Val Pro Ala Asn Phe
        755                 760                 765
Asn Glu Ala Lys Tyr Ile Ala Phe Thr Met Tyr Thr Thr Cys Ile Ile
    770                 775                 780
Trp Leu Ala Phe Val Pro Ile Tyr Phe Gly Ser Asn Tyr Lys Ile Ile
785                 790                 795                 800
Thr Met Cys Phe Ser Val Ser Leu Ser Ala Thr Val Ala Leu Gly Cys
                805                 810                 815
Met Phe Val Pro Lys Val Tyr Ile Ile Leu Ala Lys Pro Glu Arg Asn
            820                 825                 830
Val Arg Ser Ala Phe Thr Thr Ser Thr Val Val Arg Met His Val Gly
        835                 840                 845
Asp Gly Lys Ser Ser Ser Ala Ala Ser Arg Ser Ser Ser Leu Val Asn
    850                 855                 860
Leu Trp Lys Arg Arg Gly Ser Ser Gly Glu Thr Leu Arg Tyr Lys Asp
865                 870                 875                 880
Arg Arg Leu Ala Gln His Lys Ser Glu Ile Glu Cys Phe Thr Pro Lys
                885                 890                 895
Gly Ser Met Gly Asn Gly Gly Arg Ala Thr Met Ser Ser Ser Asn Gly
            900                 905                 910
Lys Ser Val Thr Trp Ala Gln Asn Glu Lys Ser Ser Arg Gly Gln His
        915                 920                 925
Leu Trp Gln Arg Leu Ser Ile His Ile Asn Lys Lys Glu Asn Pro Asn
    930                 935                 940
Gln Thr Ala Val Ile Lys Pro Phe Pro Lys Ser Thr Glu Ser Arg Gly
945                 950                 955                 960
Leu Gly Ala Gly Ala Gly Ala Gly Gly Ser Ala Gly Gly Val Gly Ala
                965                 970                 975
Thr Gly Gly Ala Gly Cys Ala Gly Ala Gly Pro Gly Gly Pro Glu Ser
            980                 985                 990
Pro Asp Ala Gly Pro Lys Ala Leu Tyr Asp Val Ala Glu Ala Glu Glu
        995                1000                1005
His Phe Pro Ala Pro Ala Arg Pro Arg Ser Pro Ser Pro Ile Ser Thr
   1010                1015                1020
Leu Ser His Arg Ala Gly Ser Ala Ser Arg Thr Asp Asp Asp Val Pro
025                1030                1035                1040
Ser Leu His Ser Glu Pro Val Ala Arg Ser Ser Ser Ser Gln Gly Ser
                1045                1050                1055
Leu Met Glu Gln Ile Ser Ser Val Val Thr Arg Phe Thr Ala Asn Ile
            1060                1065                1070
Ser Glu Leu Asn Ser Met Met Leu Ser Thr Ala Ala Pro Ser Pro Gly
        1075                1080                1085
Val Gly Ala Pro Leu Cys Ser Ser Tyr Leu Ile Pro Lys Glu Ile Gln
   1090                1095                1100
Leu Pro Thr Thr Met Thr Thr Phe Ala Glu Ile Gln Pro Leu Pro Ala
105                1110                1115                1120
Ile Glu Val Thr Gly Gly Ala Gln Pro Ala Ala Gly Ala Gln Ala Ala
                1125                1130                1135
Gly Asp Ala Ala Arg Glu Ser Pro Ala Ala Gly Pro Glu Ala Ala Ala
            1140                1145                1150
Ala Lys Pro Asp Leu Glu Glu Leu Val Ala Leu Thr Pro Pro Ser Pro
        1155                1160                1165
Phe Arg Asp Ser Val Asp Ser Gly Ser Thr Thr Pro Asn Ser Pro Val
   1170                1175                1180
Ser Glu Ser Ala Leu Cys Ile Pro Ser Ser Pro Lys Tyr Asp Thr Leu
185                1190                1195                1200
Ile Ile Arg Asp Tyr Thr Gln Ser Ser Ser Ser Leu
                1205                1210
```

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/705, G01N 33/68, C07K 16/28 — A3

(11) International Publication Number: WO 99/38975

(43) International Publication Date: 5 August 1999 (05.08.99)

(21) International Application Number: PCT/US99/02033

(22) International Filing Date: 29 January 1999 (29.01.99)

(30) Priority Data:
60/073,190 30 January 1998 (30.01.98) US

(71) Applicant *(for all designated States except US)*: UNIVERSITY TECHNOLOGY CORPORATION [US/US]; Suite 250, 3101 Iris Avenue, Boulder, CO 80301 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: JOHNSON, Thomas, E. [US/US]; 1121 West Enclave Circle, Louisville, CO 80027 (US). SIKELA, James, M. [US/US]; 6046 S. Kingston Circle, Englewood, CO 80111 (US). SIMPSON, Victoria, J. [US/US]; 1121 W. Enclave Circle, Englewood, CO 80027 (US). RIKKE, Brad, A. [US/US]; 750 Hartford, Boulder, CO 80303 (US).

(74) Agents: POLIZZI, Catherine, M. et al.; Morrison & Foerster LLP, 755 Page Mill Road, Palo Alto, CA 94304-1018 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
23 September 1999 (23.09.99)

(54) Title: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES ASSOCIATED WITH CNS DEPRESSANT SENSITIVITY AND METHODS OF USE THEREOF

(57) Abstract

The invention provides polynucleotide and polypeptide sequences associated with CNS depressant sensitivity, and methods for identifying and classifying candidate CNS depressants as well as methods to identify agents which may modulate CNS depressant action. The invention also provides mouse mGluR5 polynucleotide and polypeptide sequences.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERNATIONAL SEARCH REPORT

Inte onal Application No
PCT/US 99/02033

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/705 G01N33/68 C07K16/28

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K C12N G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 94 29449 A (SALK INST BIOTECH IND ;DAGGETT LORRIE (US); ELLIS STEVEN B (US); L) 22 December 1994 cited in the application see claim 26; example 1 ---	10-19
X	SOMMER, B. ET AL.: "RNA editing in brain controls a determinant of ion flow in glutamate-gated channels." CELL, vol. 67, 4 October 1991, pages 11-9, XP002108292 see page 16, right-hand column, paragraph 5 --- -/--	20-35

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 July 1999	27/07/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Smalt, R

2

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MINAKAMI, R. ET AL.: "The expression of two splice variants of metabotropic glutamate receptor subtype 5 in the rat brain and neural cells during development." JOURNAL OF NEUROCHEMISTRY, vol. 65, no. 4, October 1995, pages 1536-42, XP002108290 cited in the application see the whole document	20-35
X	DORRI, F. ET AL.: "Down-regulation of mGluR5 by antisense deoxynucleotides alters pharmacological responses to application of ACPD in the rat hippocampus." EXPERIMENTAL NEUROLOGY, vol. 147, no. 1, September 1997, pages 48-54, XP002108293 see abstract see page 50, left-hand column, paragraph 2	37,38
Y	MINAMI, K. ET AL.: "Effects of ethanol and anesthetics on type 1 and 5 metabotropic glutamate receptors expressed in Xenopus laevis oocytes." MOLECULAR PHARMACOLOGY, vol. 53, no. 1, 1 January 1998, pages 148-156, XP002108294 see abstract	1-9
Y	SIMPSON, V.J. ET AL.: "Mapping a major QTL for propofol neurosensitivity in F2 mice from an LS by SS cross." ANESTHESIOLOGY, vol. 65, no. 3a, September 1996, page a627 XP002108295 see the whole document	1-9
A	SCHEDL A ET AL: "A YEAST ARTIFICIAL CHROMOSOME COVERING THE TYROSINASE GENE CONFERS COPY NUMBER-DEPENDENT EXPRESSION IN TRANSGENIC MICE" NATURE, vol. 362, no. 6417, 18 March 1993, pages 258-261, XP000563891 cited in the application see the whole document	1-9

-/--

2

INTERNATIONAL SEARCH REPORT

Int ational Application No
PCT/US 99/02033

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PROCTER, M. ET AL.: "Possible role of GluR5 glutamate receptors in spinal nociceptive processing in the anaesthetized rat." JOURNAL OF PHYSIOLOGY, vol. 504P, November 1997, pages 101P-102P, XP002108296 see the whole document ---	
A	ABE T ET AL: "MOLECULAR CHARACTERIZATION OF A NOVEL METABOTROPIC GLUTAMATE RECEPTOR MGLUR5 COUPLED TO INOSITOL PHOSPHATE/CA2+ SIGNAL TRANSDUCTION" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 267, no. 19, 5 July 1992, pages 13361-13368, XP000605387 cited in the application see the whole document ---	
T	SIMPSON, V.J. ET AL.: "Ketamine exerts differential hypnotic effects in long sleep and short sleep mice." ANESTHESIOLOGY, vol. 89, no. 3a, September 1998, page A128 XP002108299 see the whole document -----	

2

INTERNATIONAL SEARCH REPORT

national application No.
PCT/US 99/02033

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found mulple (groups of) inventions in thsi international application, as follows:

1. Claims: 1-9, 20-38 completely, and 10,11,13,15,16 partially
   Mouse mGluR5 and its uses in screening and characterising CNS depressants.
2. Claims : 18 completely, and 10-13,15-17 partially
   The use of human mGluR5 in screening and characterising CNS depressants.
3. Claims; 10,11,14-16 partially
   The use of mouse mGluR1 in screening and characterigins CNS depressants.
4. Claims: 19 completely, and 10-12,14-17 partially
   The use of human mGluR1 in screening and characterising CNS depressants
5. Claims: 10-17 partially
   The use of metabotropic glutamate receptors, other than human mGluR1, human mGluR5, mouse mGlur1, and mouse mGluR5, in screening and characterising CNS depressants.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int tional Application No
PCT/US 99/02033

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9429449 A	22-12-1994	US 5521297 A	28-05-1996
		AU 685471 B	22-01-1998
		AU 7098994 A	03-01-1995
		CA 2161811 A	22-12-1994
		EP 0701611 A	20-03-1996
		GB 2286398 A,B	16-08-1995
		JP 8511168 T	26-11-1996
		US 5912122 A	15-06-1999
		US 5807689 A	15-09-1998